

FIGURE 1

CCAATCGCCCGGTGCGGTGGTGCAGGGTCTCGGGCTAGTC**ATG**GCGTCCCCGTCTCGGAGAC
TGCAGACTAAACCAGTCATTACTTGTTTCAAGAGCGTTCTGCTAATCTACACTTTTATTTTC
TGGATCACTGGCGTTATCCTTCTTGACAGTTGGCATTGTTGGGGCAAGGTGAGCCTGGAGAATTA
CTTTTCTCTTTTAAATGAGAAGGCCACCAATGTCCCCCTTCGTGCTCATTGCTACTGGTACCG
TCATTATTCTTTTGGGCACCTTTGGTTGTTTTGCTACCTGCCGAGCTTCTGCATGGATGCTA
AAACTGTATGCAATGTTTCTGACTCTCGTTTTTTTTGGTCGAACTGGTCGCTGCCATCGTAGG
ATTTGTTTTTCAGACATGAGATTAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGC
AGTATAACTCTACAGGAGATTATAGAAGCCATGCAGTAGACAAGATCCAAAATACGTTGCAT
TGTTGTGGTGTCAACGATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAAGGATT
TCCTAAGAGTTGCTGTAAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACA
ATGAAGGTTGTTTTATAAAGGTGATGACCATTATAGAGTCAGAAATGGGAGTCGTTGCAGGA
ATTTCCTTTGGAGTTGCTTGCTTCCAACGATTGGAATCTTTCTCGCCTACTGCCWCTCTCG
TGCCATAACAAATAACCAGTATGAGATAGTG**TAA**CCCAATGTATCTGTGGGCCTATTCCTCT
CTACCTTTAAGGACATTTAGGGTCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAACTG
ACAACACTACTTACTGATAGACCAAAAAACTACACCAGTAGGTTGATTCAATCAAGATGTAT
GTAGACCTAAAACTACACCAATAGGCTGATTCAATCAAGATCCGTGCTCGCAGTGGGCTGAT
TCAATCAAGATGTATGTTTGCTATGTTCTAAGTCCACCTTCTATCCCATTTCATGTTAGATCG
TTGAAACCCTGTATCCCTCTGAAACACTGGAAGAGCTAGTAAATTGTAAATGAAGT

FIGURE 2

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA19902

><subunit 1 of 1, 245 aa, 1 stop, 1 unknown

><MW: -1, pI: 8.36, NX(S/T): 1

MASPSRRLQTKPVITCFKSVLLIYTFIFWITGVILLAVGIWGVSLNYSLLNEKATNVPF
VLIATGTVIILLGTFGCFATCRASAWMLKLYAMFLTLVFLVELVAAIVGFVFRHEIKNSFKN
NYEKALKQYNSTGDYRSHAVDKIQNTLHCCGVTDYRDWTDNYYSEKGFPSCKLEDCTPQ
RDADKVNNEGCFIKVMTIIESEMGVVAGISFGVACFQLIGIFLAYCXSRITNNQYEIV

Important features of the protein:

Signal peptide:

amino acids 1-42

Transmembrane domains:

amino acids 19-42, 61-83, 92-114, 209-230,

N-glycosylation site.

amino acids 134-138

Tyrosine kinase phosphorylation site.

amino acids 160-168, 160-169

N-myristoylation site.

amino acids 75-81, 78-84, 210-216, 214-220, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 69-80, 211-222

FIGURE 3

CCCACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCACAAA
GGAGTCGCGCCGCGCCGCGCCGCCCCCTCCCTCCGGTGGGCCCCGGGAGGTAGAGAAAGTCAGT
GCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT
TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGG
GCCGGGGTAGGCTCTGGAAAGGGCCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA
GCCGAGAGGTTTTCCACCGAGGCCCCGCGCTTGAGGGATCTGAAGAGGTTCCCTAGAAGAGGGT
GTTCCCTCTTTGGGGGTCCTCACCAGAAGAGGTTCTTGGGGGTCGCCCTTCTGAGGAGGCT
GCGGCTAACAGGGCCCAGAACTGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG
GGAATAAGCTCTGCAACTTTCTTTGGCATTGAGTTGTTAAAAACAAATAGGATGCAAATTCC
TCAACTCCAGGTTATGAAAACAGTACTTGAAAACTGAAAACTACCTAAATGATCGTCTTTG
GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGCCAGGGTCTGTTGTTGACTCTCGAAGAG
CACATAGCCCACTTCCTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATTCCCTGTATCTG
CCGAGATGACAGTGGAACAGATGACAGTGTTGACACCCAACAGCAACAGGCCGAGAACAGTG
CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCTGTTCGGCCACCAAGGAGGGGC
CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTTGGACACACT
GGCAGTAATACGGACTCTTGTAGATAAGTAAGTATCTGACTCACGGTCACCTCCAGTGGAAT
GAAAGTGTTCTGCCCCGAACCATGACTTTAGGACTCCTTCAGTTCCTTTAGGACATACTCG
CCAAGCCTTGTGCTCACAGGGCAAAGGAGAATATTTTAATGCTCCGCTGATGGCAGAGTAAA
TGATAAGATTTGATGTTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC
GTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

FIGURE 4

MIVFGWAVFLASRSLGQGLLLTLEEHIAHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQQQ
AENSAVPTADTRSQPRDPVRPPRRGRGPHEPRRKKQNV DGLVLDTLAVIRTLVDKO

Signal peptide:

amino acids 1-16

Casein kinase II phosphorylation site.

amino acids 22-26, 50-54, 113-117

N-myristoylation site.

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

FIGURE 5

GGCACGAGGCGCTGTCCACCCGGGGGCGTGGGAGTGAGGTACCAGATTCAGCCCATTTGGCC
CCGACGCCTCTGTTCTCGGAATCCGGGTGCTGCGGATTGAGGTCCCGGTTCCTAACGGACTG
CAAG**ATG**GAGGAAGGCGGGAACCTAGGAGGCCTGATTAAGATGGTCCATCTACTGGTCTTGT
CAGGTGCCTGGGGCATGCAAATGTGGGTGACCTTCGTCTCAGGCTTCCTGCTTTTCCGAAGC
CTTCCCCGACATACCTTCGGACTAGTGCAGAGCAAACCTTCCCCTTCTACTTCCACATCTC
CATGGGCTGTGCCTTCATCAACCTCTGCATCTTGGCTTCACAGCATGCTTGGGCTCAGCTCA
CATTCTGGGAGGCCAGCCAGCTTTACCTGCTGTTCCCTGAGCCTTACGCTGGCCACTGTCAAC
GCCCCTGGCTGGAACCCCGCACCACAGCTGCCATGTGGGCCCTGCAAACCGTGGAGAAGGA
GCGAGGCCTGGGTGGGGAGGTACCAGGCAGCCACCAGGGTCCCGATCCCTACCGCCAGCTGC
GAGAGAAGGACCCCAAGTACAGTGCTCTCCGCCAGAATTTCTTCCGCTACCATGGGCTGTCC
TCTCTTTGCAATCTGGGCTGCGTCCTGAGCAATGGGCTCTGTCTCGCTGGCCTTGCCCTGGA
AATAAGGAGCCTC**TAG**CATGGGCCCTGCATGCTAATAAATGCTTCTTCAGAAATGAAAAAA
AAAAAAAAAAAA

FIGURE 6

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56107

<subunit 1 of 1, 231 aa, 1 stop

<NX(S/T): 0

MEEGGN LGGLIKMVHLLVLSGAWGMQMWVTFVSGFLLFRSLPRHTFGLVQSKLFPFYFHISM
GCAFINLCILASQHAWAQLTFWEASQLYLLFLSLTLATVNARWLEPRTTAAMWALQTVEKER
GLGGEVPGSHQGPDPYRQLREKDPKYSALRQNFFRYHGLSSLCNLGCVLSNGLCLAGLALEIRSL

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 86-103, 60-75

Casein kinase II phosphorylation site.

amino acids 82-86

Tyrosine kinase phosphorylation site.

amino acids 144-151

N-myristoylation site.

amino acids 4-10, 5-11, 47-53, 170-176, 176-182

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 54-65

G-protein coupled receptors proteins.

amino acids 44-85

FIGURE 7

AATTCAGATTTTAAAGCCCATTCTGCAGTGGAATTTTCATGAACTAGCAAGAGGACACCATCTT
CTTGTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAAATGCTCTTTTGGGT
GCTAGGCCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAACTAAAGATTGAAG
ACATCACTGATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAAACTTGGCAGCC
AGAACTTTTGATAAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAAC
AGCTTTAAAGGCAGAAACCTCAGAGAGACTTCGTA CTGTGCTTCTGGATGTGACCGACCCAG
AGAATGTCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGT
CTGATCAATAATGCTGGTGTTC CCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGA
CTACAGAGAACCTATTGAAGTGAACCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTC
CTTTGGTCAAGAAAGCTCAAGGGAGAGTTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCA
ATCGTTGGAGGGGGCTATACTCCATCCAAATATGCAGTGGAAGGTTTCAATGACAGCTTAAG
ACGGGACATGAAAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAA
ACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCA
GACATCAAACAACAATATGGAGAAGGTTACATTGAAAAAAGTCTAGACAAACTGAAAGGCAA
TAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTGCATGGACCACGCTCTAACAA
GTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAATTTTCTGGATACCTCTG
TCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAA
TCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGAT
TTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACTCATTTAGA
TCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGTCCCTG
CTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCTGT
ATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATG
ATCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTT
AAGTATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 8

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56406

><subunit 1 of 1, 319 aa, 1 stop

><MW: 35227, pI: 8.97, NX(S/T): 3

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKG FHVIAACLT
 ESGSTALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDW
 LTLEDYREPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGF
 NDSLRRDMKAFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLD
 KLKGNKSYVNMDLSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQK
 AELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 9

GCGGGCTGTTGACGGCGCTGCG**ATG**GCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGGTT
 CCTCTCAGTCGGACTTCCTGACGCCGCCAGTGGGCGGGGCCCTTGGGCCGTCGCCACCACT
 GTAGTCATGTACCCACCGCCGCCGCCGCCCTCATCGGGACTTCATCTCGGTGACGCTGAG
 CTTTGGCGAGAGCTATGACAACAGCAAGAGTTGGCGGGCGGCTCGTGCTGGAGGAAATGGA
 AGCAACTGTCGAGATTGCAGCGGAATATGATTCTCTTCCTCCTTGCCTTTCTGCTTTTCTGT
 GGACTCCTCTTCTACATCAACTTGGCTGACCATTTGGAAAGCTCTGGCTTTCAGGCTAGAGGA
 AGAGCAGAAGATGAGGCCAGAAATTGCTGGGTAAAACCAGCAAATCCACCCGTCTTACCAG
 CTCCTCAGAAGGCGGACACCGACCCTGAGAACTTACCTGAGATTTTCGTACAGAAGACACAA
 AGACACATCCAGCGGGGACCACCTCACCTGCAGATTAGACCCCCAAGCCAAGACCTGAAGGA
 TGGGACCCAGGAGGAGGCCACAAAAGGCAAGAAGCCCCCTGTGGATCCCCGCCCGGAAGGAG
 ATCCGCAGAGGACAGTCATCAGCTGGAGGGGAGCGGTGATCGAGCCTGAGCAGGGCACCAGAG
 CTCCCTTCAAGAAGAGCAGAAGTGCCACCAAGCCTCCCCTGCCACCGGCCAGGACACAGGG
 CACACCAGTGCATCTGAACTATCGCCAGAAGGGCGTGATTGACGTCTTCCTGCATGCATGGA
 AAGGATACCGCAAGTTTGCATGGGGCCATGACGAGCTGAAGCCTGTGTCCAGGTCTTCAGT
 GAGTGGTTTGGCCTCGGTCTCACACTGATCGACGCGCTGGACACCATGTGGATCTTGGGTCT
 GAGGAAAGAATTTGAGGAAGCCAGGAAGTGGGTGTGGAAGAAGTTACACTTTGAAAAGGACG
 TGGACGTCAACCTGTTTGAAGAGCAGATCCGCATCCTGGGGGGGCTCCTGAGTGCCTACCAC
 CTCTCTGGGGACAGCCTCTTCCTGAGGAAAGCTGAGGATTTTGGAAATCGGCTAATGCCTGC
 CTTGAGAACACCATCCAAGATTCCCTTACTCGGATGTGAACATCGGTACTGGAGTTGCCACCC
 CGGCACGGTGGACCTCCGACAGCACTGTGGCCGAGGTGACCAGCATTAGCTGGAGTTCCGG
 GAGCTCTCCCGTCTCACAGGGGATAAGAAGTTTCAGGAGGCAGTGGAGAAGGTGACACAGCA
 CATGCACGGCCTGTCTGGGAAGAAGGATGGGCTGGTGCCCATGTTTCATCAATACCCACAGTG
 GCGTCTTACCCACCTGGGCGTATTACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTAC
 CTGCTGAAGCAGTGGATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGA
 AGCTATCGAGGGTGTGAGAACGCACCTGCTGCGGCACTCCGAGCCCAGTAAGCTCACCTTTG
 TGGGGAGCTTGCCACGCGCGCTTCAGTGCCAAGATGGACCACCTGGTGTGCTTCCTGCCA
 GGGACGCTGGCTCTGGGCGTCTACCACGGCCTGCCCGCCAGCCACATGGAGCTGGCCCAGGA
 GCTCATGGAGACTTGTTACCAGATGAACCGGCAGATGGAGACGGGGCTGAGTCCCGAGATCG
 TGGACTTCAACCTTTACCCCCAGCCGGGCCGTGCGGACGTGGAGGTCAAGCCAGCAGACAGG
 CACAACCTGCTGCGGCCAGAGACCGTGGAGAGCCTGTTCTACCTGTACCGCGTCACAGGGGA
 CCGCAAATACCAGGACTGGGGCTGGGAGATTCTGCAGAGCTTCAGCCGATTCACACGGGTCC
 CCTCGGGTGGCTATTCTTCCATCAACAATGTCCAGGATCCTCAGAAGCCCCGAGCCTAGGGAC
 AAGATGGAGAGCTTCTTCCTGGGGGAGACGCTCAAGTATCTGTTCTTGCTCTTCTCCGATGA
 CCCAAACCTGCTCAGCCTGGACGCCTACGTGTTCAACACCGAAGCCCACCCTCTGCCTATCT
 GGACCCCTGCC**TAG**GGTGGATGGCTGCTGGTGTGGGGACTTCGGGTGGGCAGAGGCACCTTG
 CTGGGTCTGTGGCATTTCCTCAAGGGCCACGTAGCACCGGCAACCGCCAAGTGGCCCAGGCT
 CTGAACTGGCTCTGGGCTCCTCCTCGTCTCTGCTTTAATCAGGACACCGTGAGGACAAGTGA
 GGCCGTCACTTGGTGTGATGCGGGGTGGGCTGGGCCGCTGGAGCCTCCGCTGCTTCCTC
 CAGAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAGGTCTCTGTGGGCCGACCA
 GAGGGGGGCTTCGAGGTGGTCCCTGGTACTGGGGTGACCGAGTGGACAGCCCAGGGTGACG
 TCTGCCCGGGCTCGTGAAGCCTCAGATGTCCCAATCCAAGGGTCTGGAGGGGCTGCCGTGA
 CTCCAGAGGCCTGAGGCTCCAGGGCTGGCTCTGGTGTGTTTACAAGCTGGACTCAGGGATCCTC
 CTGGCCGCCCCGACAGGGGGCTTGAGGGCTGGACGGCAAGTCCGTCTAGCTCACGGGCCCCCT
 CCAGTGGAATGGGTCTTTTCGGTGGAGATAAAAGTTGATTGCTCTAACCGCAA

FIGURE 10

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56529

><subunit 1 of 1, 699 aa, 1 stop

><MW: 79553, pI: 7.83, NX(S/T): 0

MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVMYPPPPPPPHRDFISVTLSFGESYDN
 SKSWRRRSCWRKWKQLSRLQRMILFLLAFLLFCGLLFYINLADHWKALAFRLEEEQKMRPE
 IAGLKPANPPVLPAPQKADTDPENLPEISSQKTQRHIQRGPPHLQIRPPSQDLKDGTQEEAT
 KRQEAPVDPRPEGDPQRTVISWRGAVIEPEQGTELPSRRAEVPTKPPLPPARTQGTPVHLNY
 RQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSFSEWFGLGLTLIDALDTMWILGLRKEFEEA
 RKWVSKKLHFEKDVDVNLFEISTIRILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKI
 PYSDVNIGTGVAHPPRWTSDSTVAEVTSIQLEFREL SRLTGDKKFQEAVEKVTQHIHGLSGK
 KDGLVPMFINTHSGLFTHLGVFTLGARADSYEYLLKQWIQGGKQETQLLEDYVEAIEGVRT
 HLLRHSEPSKLTFFVGE LAHGRFSAKMDHLVCFLPGTLALGVYHGLPASHMELAQELMETCYQ
 MNRQMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLRLPETVESLFYLYRVTGDRKYQDWG
 WEL LQSFSRFTRVPSSGGYSSINNVQDPQKPEPRDKMESFFLGETLKYLFLLFSDDPNLLSLD
 AYVFNTEAHPLPIWTPA

Important features of the protein:

Transmembrane domain:

amino acids 21-40 and 84-105 (type II)

FIGURE 11

GGCGCCGCGTAGGCCCCGGGAGGCCGGGCCGGCCGGGCTGCGAGCGCCTGCCCCATGCGCCGC
CGCCTCTCCGCACG**ATG**TTCCCCCTCGCGGAGGAAAGCGGCGCAGCTGCCCTGGGAGGACGGC
AGGTCCGGGTTGCTCTCCGGCGGCCTCCCTCGGAAGTGTTCCGTCTTCCACCTGTTTCGTGGC
CTGCCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGGACGTGG
CCCGGGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCCGTGCCTGCCCCCAGAG
CCGCCCCCTGAGCACTGGGAAGAAGACGCATCCTGGGGCCCCCACCGCCTGGCAGTGCTGGT
GCCCTTCCGCGAACGCTTCGAGGAGCTCCTGGTCTTCGTGCCCCACATGCGCCGCTTCCTGA
GCAGGAAGAAGATCCGGCACCACATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTCAAC
CGGGCAGCGCTCATCAACGTGGGCTTCCTGGAGAGCAGCAACAGCACGGACTACATTGCCAT
GCACGACGTTGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTTCCTGAGGCTGGGC
GCTTCCACGTGGCCTCCCCGGAGCTCCACCCTCTCTACCACTACAAGACCTATGTCGGCGGC
ATCCTGCTGCTCTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGG
CTGGGGCCGCGAGGACGACGAGTTCTACCGGCGCATTAAGGGAGCTGGGCTCCAGCTTTTCC
GCCCTCGGGAATCACAACCTGGGTACAAGACATTTGCCACCTGCATGACCCAGCCTGGCGG
AAGAGGGACCAGAAGCGCATCGCAGCTCAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGG
AGGCCTGAACACTGTGAAGTACCATGTGGCTTCCCGCACTGCCCTGTCTGTGGGCGGGGCC
CCTGCACTGTCTCAACATCATGTTGGACTGTGACAAGACCGCCACACCCTGGTGCACATTC
AG**CTGA**GCTGGATGGACAGTGAGGAAGCCTGTACCTACAGGCCATATTGCTCAGGCTCAGGA
CAAGGCCTCAGGTCGTGGGCCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCA
AGCTACGCAATTGCAGCCACCCGGCCGCCAAGGCAGGCTTGGGCTGGGCCAGGACACGTGGG
GTGCCTGGGACGCTGCTTGCCATGCACAGTGATCAGAGAGAGGCTGGGGTGTGTCCTGTCCG
GGACCCCCCTGCCTTCCTGCTCACCTACTCTGACCTCCTTCACGTGCCAGGCCTGTGGG
TAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCTACTCTGACCTCCTTCACGTGCCC
AGGCCTGTGGGTAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCCCCAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 12

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531

><subunit 1 of 1, 327 aa, 1 stop

><MW: 37406, pI: 9.30, NX(S/T): 1

MFPSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHLFVACL~~SL~~GFFSLLWLQLSCSGDVARAVR
 GQGQETSGPPRACPPEPPPEHWEEDASWGPHRLAVLVPPFRERFEELLVFVPHMRRFLSRKKI
 RHHIYVLNQVDHFRFNRAALINVGFL~~ESS~~NSTDYIAMHDVDLLPLNEELDYGFP~~EAG~~PFHVA
 SPELHPLYHYKTYVGGILLLSKQHYRLCNGMSNRFWGWGREDDEFYRRIKGAGLQLFRPSGI
 TTGYKTFRHLHDP~~AWR~~KRDQKRIAAQKQEQFKVDREGGLNTVKYHVASRTALSVGGAPCTVL
 NIMLDCDKTATPWCTFS

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 29-49 (type II)

N-glycosylation site.

amino acids 154-158

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 27-31

Tyrosine kinase phosphorylation site.

amino acids 226-233

N-myristoylation site.

amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310

FIGURE 13

CAATGTTTGCCTATCCACCTCCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCT
GCTGCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCT
CTTCTGCCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTC
CTGCCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGT
GGGGGATGGCTAAGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGG
GCTGCCTTATTTAAAGTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGC
CCTGTTCATTAAGAAATTGTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGT
AAATATGTCTTTATAATAAACAGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

13/249

FIGURE 14

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56862

<subunit 1 of 1, 73 aa, 1 stop

<MW: 7879, pI: 7.21, NX(S/T): 0

MLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFPPSFLCLLPHRPAMTCSQAQ
PRGEGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 15

GGGACCCATGCGGCCGTGACCCCCGGCTCCCTAGAGGCCAGCGCAGCCGCAGCGGACAAAG
 GAGCATGTCCGCGCCGGGGAAGGCCCGTCTCCGGCCGCCATAAGGCTCCGGTCGCCGCTGG
 GCGCGCGCCGCGCTCCTGCCCCCGGGCTCCGGGGCGGCCGCTAGGCCAGTGCGCCGCCG
 CTCGCCCCGCAGGCCCGGGCCGCGAG**CATG**GAGCCACCCGGACGCCGGCGGGGCCGCGCGCA
 GCCGCCGCTGTTGCTGCCGCTCTCGCTGTTAGCGCTGCTCGCGCTGCTGGGAGGCGGCGGCG
 GCGGCGGCGCCGCGGCGCTGCCCGCCGGCTGCAAGCACGATGGGCGGCCCGAGGGGCTGGC
 AGGGCGGCGGGCGCCGCCGAGGGCAAGGTGGTGTGCAGCAGCCTGGAACCTCGCGCAGGTCTT
 GCCCCAGATACTCTGCCCAACCGCACGGTCACCCCTGATTCTGAGTAACAATAAGATATCCG
 AGCTGAAGAATGGCTCATTTTCTGGGTAAAGTCTCCTTGAAAGATTGGACCTCCGAAACAAT
 CTTATTAGTAGTATAGATCCAGGTGCCTTCTGGGGACTGTCATCTCTAAAAAGATTGGATCT
 GACAAACAATCGAATAGGATGTCTGAATGCAGACATATTTGAGGACTCACCAATCTGGTTC
 GGCTAAACCTTTCGGGGAATTTGTTTTCTTCATTATCTCAAGGAACTTTTGATTATCTTGCG
 TCATTACGGTCTTTGGAATTCCAGACTGAGTATCTTTTGTGTGACTGTAACATACTGTGGAT
 GCATCGCTGGGTAAAGGAGAAGAACATCACGGTACGGGATACCAGGTGTGTTTATCCTAAGT
 CACTGCAGGCCCAACCAGTCACAGGCGTGAAGCAGGAGCTGTTGACATGCGACCCCTCCGCTT
 GAATTGCCGCTCTTCTACATGACTCCATCTCATCGCCAAGTTGTGTTTGAAGGAGACAGCCT
 TCCTTTCCAGTGCATGGCTTCATATATTGATCAGGACATGCAAGTGTGTTGTTGATCAGGATG
 GGAGAATAGTTGAAACCGATGAATCGCAAGGTATTTTTGTTGAAAAGAACATGATTCAACAAC
 TGCTCCTTGATTGCAAGTGCCCTAACCATTTCTAATATTCAGGCTGGATCTACTGGAATTG
 GGGCTGTGATGTCCAGACCAACGTGGGAATAATACGAGGACTGTGGATATTGTGGTATTAG
 AGAGTTCTGCACAGTACTGTCCTCCAGAGAGGGTGGTAAACAACAAAGGTGACTTCAGATGG
 CCCAGAACATTGGCAGGCATTACTGCATATCTGCAGTGTACGCGGAACACCCATGGCAGTGG
 GATATATCCCGGAAACCCACAGGATGAGAGAAAAGCTTGGCGCAGATGTGATAGAGGTGGCT
 TTGGGCAGATGATGATTATTCTCGCTGTGAGTATGCAATGATGTCACTAGAGTTCTTTAT
 ATGTTTAATCAGATGCCCTCAATCTTACCAATGCCGTGGCAACAGCTCGACAGTTACTGGC
 TTACACTGTGGAAGCAGCCAACCTTTCTGACAAAATGGATGTTATATTTGTGGCAGAAATGA
 TTGAAAAATTTGGAAGATTTACCAAGGAGGAAAAATCAAAGAGCTAGGTGACGTGATGGTT
 GACATTGCAAGTAACATCATGTTGGCTGATGAACGTGTCTGTGGCTGGCGCAGAGGGAAGC
 TAAAGCCTGCAGTAGGATTGTGCAGTGTCTTCAGCGCATTGCTACCTACCGGCTAGCCGGTG
 GAGCTCACGTTTATTCAACATATTACCCCAATATTGCTCTGGAAGCTTATGTCATCAAGTCT
 ACTGGCTTCACGGGGATGACCTGTACCGTGTTCAGAAAGTGGCAGCCTCTGATCGTACAGG
 ACTTTCCGATTATGGGAGGCGGGATCCAGAGGGGAAACCTGGATAAGCAGCTGAGCTTTAAGT
 GCAATGTTTCAAATACATTTTCGAGTCTGGCACTAAAGGTATGTTACATTCTGCAATCATTT
 AAGACTATTTACAGT**TAA**ATTAGAATGCTCCAAATGTTCTGCTTCGCAAAATAACCTTATTA
 AAAGATTTTTTTTTTGCAAGGAAGATAGGTATTATTGCTTTTGCTACTGTTTTAAAGAAAATA
 ACCAGGAAGAACTGCATTACGACTTTCAAGGGCCCTAGGCATTTTTGCCTTTGATTCCCTTT
 CTTACATAAAAAATATCAGAAATTACATTTTATAACTGCAGTGGTATAAATGCAAATATACT
 ATTGTTACATGTGAAAAAATTTTATTTGACTTAAAAGTTTATTTATTTGTTTTTTTGCTCCT
 GATTTTAAGACAATAAGATGTTTTTCATGGGCCCTAAAAGTATCATGAGCCTTTGGCACTGC
 GCCTGCCAAGCCTAGTGGAGAAGTCAACCCTGAGACCAGGTGTTTAATCAAGCAAGCTGTAT
 ATCAAAATTTTTGGCAGAAAACACAAATATGTCATATATCTTTTTTTAAAAAAGTATTTCA
 TTGAAGCAAGCAAAATGAAAGCATTTTTACTGATTTTTAAATTTGGTGCTTTTAGATATATTT
 GACTACACTGTATTGAAGCAAAATAGAGGAGGCACAACTCCAGCACCTAATGGAACCACATT
 TTTTTCACCTTAGCTTTCTGTGGGCATGTGTAATTGTATTCTCTGCGGTTTTTAATCTCACAG
 TACTTTATTTCTGTCTTGTCCCTCAATAATATCACAAACAATATTCCAGTCATTTTAATGGC
 TGCATAATAACTGATCCAACAGGTGTTAGGTGTTCTGGTTTAGTGTGAGCACTCAATAAATA
 TTGAATGAATGAACGAAAAAAAAAAAAAAAAA

FIGURE 16

MEPPGRRRRGRAQPPLLLPLSLLALLLGGGGGGGAAALPAGCKHDGRPRGAGRAAGAAEGK
 VVCSSLELAQVLPPDTLPNRTVTLILSNNKISELKNGSFSGLSLLERLDLRNNLISSIDPGA
 FWGLSSLKRLDLTNNRIGCLNADIFRGLTNLVRLNLSGNLFSSLSQGTFDYLASLRSLFQT
 EYLLCDCNILWMHRWVKEKNITVRDTRCVYPKSLQAQPVTVGVKQELLTCDPPELPSFYMTP
 SHRQVVFEGDSLPPFQCMASYIDQDMQVLWYQDGRIVETDESQGI FVEKNMIHNCSLIASALT
 ISNIQAGSTGNWVGCHVQTKRGNNTRTVDIVVLESSAQYCPPERVVNNKGDFRWPRTLGIT
 YLQCTRNTHGSGIYPGNPQDERKAWRRCDRGGFWADDDYSRCQYANDVTRVLYMFNQMPNL
 TNAVATARQLLAYTVEAANFSDKMDVIFVAEMIEKFGRFTKEEKSKELGDMVDIASNIMLA
 DERVLWLAQREAKACSRIVQCLQRIATYRLAGGAHVYSTYSPNIALEAYVIKSTGFTGMTCT
 VFQKVAASDRTGLSDYGRRDPEGNLDKQLSFKCNVSNTFSSLALKVCYILQSFKTIYS

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 13-40 (type II)

N-glycosylation site.

amino acids 81-85, 98-102, 159-163, 206-210, 301-305, 332-336,
 433-437, 453-457, 592-596

N-myristoylation site.

amino acids 29-35, 30-36, 31-37, 32-38, 33-39, 34-40, 51-57,
 57-63, 99-105, 123-129, 142-148, 162-168, 317-323, 320-326,
 384-390, 403-409, 554-560

FIGURE 17

CCGTGGGG**ATG**TCTAGGAGCTCGAAGGTGGTGCTGGGCCTCTCGGTGCTGCTGACGGCGGCC
ACAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGCTTCGTGACGGAGTTAT
CAGAGACATTGAGAGGCCAAATTCGGAAAAAAGAAAACATTCGTCTTTTGGGAGAACAGATTA
TTTTGACTGAGCAACTTGAAGCAGAAAGAGAGAAGATGTTATTGGCAAAAGGATCTCAAAAA
TCAT**TGA**CTTGAATGTGAAATATCTGTTGGACAGACAACACGAGTTTGTGTGTGTGTGTTGAT
GGAGAGTAGCTTAGTAGTATCTTCATCTTTTTTTTTTGGTCACTGTCCTTTTAAACTTGATCA
AATAAAGGACAGTGGGTCATATAAGTTACTGCTTTCAGGGTCCCTTATATCTGAATAAAGGA
GTGTGGGCAGACACTTTTTTGAAGAGTCTGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGG
TCACTGTCCAGTGCTTAGGGTTGTTACTGAGAAGCACTGCCGAGCTTGTGAGAAGGAAGGGA
TGGATAGTAGCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAACA
TCGACCTCAGAAGGACTGGAGGAAGGTGAAGTGGAGGGAGAGACGCTCCTGATCGTCGAATCC

FIGURE 18

MSRSSKVVLGLSVLLTAATVAGVHVKQQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILT
EQLEAEREKMLLAKGSQKS

Signal peptide:

amino acids 1-21

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 19

CTGTCGTCTTTGCTTCAGCCGAGTCGCCACTGGCTGCCTGAGGTGCTCTTACAGCCTGTTC
 CAAGTGTGGCTTAATCCGTCTCCACCACCAGATCTTTCTCCGTGGATTCTCTGCTAAGACC
 GCTGCCATGCCAGTGACGGTAACCCGCACCACCATCACAAACCACCACGACGTATCTTCGGG
 CCTGGGGTCCCCCATGATCGTGGGGTCCCCTCGGGCCCTGACACAGCCCCTGGGTCTCCTTCGC
 CTGCTGCAGCTGGTGTCTACCTGCGTGGCCTTCTCGCTGGTGGCTAGCGTGGGCGCCTGGAC
 GGGGTCCATGGGCAACTGGTCCATGTTACCTGGTGCTTCTGCTTCTCCGTGACCCTGATCA
 TCCTCATCGTGGAGCTGTGCGGGCTCCAGGCCCGCTTCCCCCTGTCTTGGCGCAACTTCCCC
 ATCACCTTCGCCTGCTATGCGGGCCCTCTTCTGCCTCTCGGCCCTCCATCATCTACCCACCAC
 CTATGTCCAGTTCTGTCCCACGGCCGTTCGCGGGACCACGCCATCGCCGCCACCTTCTTCT
 CCTGCATCGCGTGTGTGGCTTACGCCACCGAAGTGGCCTGGACCCGGGGCCCGGGCCGGCGAG
 ATCACTGGCTATATGGCCACCGTACCCGGGGCTGCTGAAGGTGCTGGAGACCTTCGTTGCCTG
 CATCATCTTCGCGTTCATCAGCGACCCCAACCTGTACCAGCACCAGCCGGCCCTGGAGTGGT
 GCGTGGCGGTGTACGCCATCTGCTTCATCCTAGCGGCCATCGCCATCCTGCTGAACCTGGGG
 GAGTGCACCAACGTGCTACCCATCCCCTTCCCCAGCTTCTGTGCGGGCTGGCCTTGCTGTC
 TGTCTCTCTATGCCACCGCCCTTGTCTCTGGCCCCCTTACCAGTTCGATGAGAAGTATG
 GCGGCCAGCCTCGGCGCTCGAGAGATGTAAGCTGCAGCCGCAGCCATGCCTACTACGTGTGT
 GCCTGGGACCGCCGACTGGCTGTGGCCATCCTGACGGCCATCAACCTACTGGCGTATGTGGC
 TGACCTGGTGCACCTCTGCCCACCTGGTTTTTGTCAAGGTCTTAAGACTCTCCCAAGAGGCTCC
 GGTTCCTCTCCAACCTCTTTGTTCTTCTTGGCCGAGTTTTCTTTATGGAGTACTTCTTTCC
 TCCGCCTTTCCTCTGTTTTCTCTTCTTCTGCTCTCCCCTCCCTCCACCTTTTTCTTCTTCC
 CAATTCCTTGCACTCTAACCAGTTCTTGGATGCATCTTCTTCCCTTCCCTTTCCTCTTGCTGT
 TTCCTTCCCTGTGTTGTTTTGTTGCCACATCCTGTTTTACCCCTGAGCTGTTTCTCTTTTT
 CTTTTCTTTCTTTTTTTTTTTTTTTTTTAAGACGGATTCTCACTCTGTGGCCCAGGCTGGAG
 TGCAGTGGTGCATCTCAGCTCACTGCAACCCCCGCCCTCCTGGGTTCAAGCGATTCTCCTCC
 CCCAGCCTCCCAAGTAGCTGGGAGGACAGGTGTGAGCTGCCGCACCCAGCCTGTTTCTCTTT
 TTCCACTCTTCTTTTTTCTCATCTCTTTTCTGGGTTGCCTGTTCGGCTTTCTTATCTGCCTGT
 TTTGCAAGCACCTTCTCCTGTGTCCTTGGGAGCCCTGAGACTTCTTTCTCTCCTTGCCTCCA
 CCCACCTCCAAAGGTGCTGAGCTCACATCCACACCCCTTGCAGCCGTCCATGCCACAGCCCC
 CCAAGGGGCCCCATTGCCAAAGCATGCCTGCCACCCCTCGCTGTGCCTTAGTCAGTGTGTAC
 GTGTGTGTGTGTGTGTGTTGGGGGGTGGGGGGTGGGTAGCTGGGGATTGGGCCCTCTTTCT
 CCCAGTGGAGGAAGGTGTGCAGTGTACTTCCCCTTTAAATTAAAAAACATATATATATATAT
 ATTTGGAGGTGAGTAATTTCCAATGGGCGGGAGGCATTAAAGCACCGACCCTGGGTCCCTAGG
 CCCCCCTGGCACTCAGCCTTGCCAGAGATTGGCTCCAGAATTTTTGCCAGGCTTACAGAACAC
 CCACTGCCTAGAGGCCATCTTAAAGGAAGCAGGGGGCTGGATGCCTTTCATCCCAACTATTCT
 CTGTGGTATGAAAAAG

FIGURE 20

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58727

<subunit 1 of 1, 322 aa, 1 stop

<MW: 35274, pI: 8.57, NX(S/T): 1

MPVTVTRTTITTTTTSSSGLGSPMIVGSPRALTQPLGLLRLLQLVSTCVAFSLVASVGAWTG
SMGNWSMFTWCFCFSVTLIILIVELCGLQARFPLSWRNFPI TFACYAALFCLSASIIYPTTY
VQFLSHGRSRDHAI AATFFSCIACVAYATEVAWTRARPGEITGYMATVPGLLKVLETFVACI
IFAFISDPNLYQHQP ALEWCVAVYAICFILAAIAILLNLGECTNVLPIPFPSFLSGLALLSV
LLYATALVLWPLYQFDEKYGGQPRRSRDVSCSRSHAYYVCAWDRRLAVAILTAINLLAYVAD
LVHSAHLVFVKV

Important features:**Transmembrane domains:**

amino acids 41-60 (type II), 66-85, 101-120, 137-153, 171-192,
205-226, 235-255 and 294-312

N-glycosylation site.

amino acids 66-69

Glycosaminoglycan attachment site.

amino acids 18-21

[illegible]

FIGURE 22

MFLATLSFLLPFAHPFGTVSCEYMLGSPSSLAQVNLSPFSPKVMHMDPNYCHPSTSLHLC
LAWSFTRLLHPPLSPGISQVVKDHSVTKPTAMAQGRVAHLIEWKGWSKPSDSPAALES
AFSSYSDLSEGEQEARFAAGVAEQFAIAEAKLRAWSSVDGEDSTDDSYDEDFAGGMDT
DMAGQLPLGPHLQDLFTGHRFSRPVRQGSVEPESDCSQTVPDTLCSSLC
SLEDGLLGSPARLASQLLGDELLAKLPPSRESAFRSLGPLEAQDSL
YNSPLTESCLSPAEEEPAPCKDCQPLCPPLTG
SWERQRQASDLASSGVVSLDEDEAEPEEQ

Signal peptide:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 123-127, 128-132, 155-159, 162-166, 166-170, 228-232,
285-289, 324-328

Tyrosine kinase phosphorylation site.

amino acids 44-52

N-myristoylation site.

amino acids 17-23, 26-32, 173-179

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 11-22

FIGURE 23

GGTTCCTGGGCGCTCTGTTACACAAGCAAGATACAGCCAGCCCCACCTAATTTTGTTCCT
GGCACCTCCTGCTCAGTGCACATTGTACACTTAACCCATCTGTTTTCTCTAATGCACGA
CAGATTCCTTTCAGACAGGACAACCTGTGATATTTAGTTTCCTGATTGTAAATACCTCCTAAG
CCTGAAGCTTCTGTTACTAGCCATTGTGAGCTTCAGTTTCTTCATCTGCAAAATGGGCATAA
TACAATCTATTCTTGCCACATCAAGGGATTGTTATTCCTTTAAAAAAAACCAATACCAAAG
AAGCCTACAATGTTGGCCTTAGCCAAAATTCTGTTGATTTCAACGTTGTTTTATTCACTTCT
ATCGGGGAGCCATGGAAAAGAAAATCAAGACATAAACACAACACAGAACATTGCAGAAGTTT
TTAAACAATGGAAAATAAACCTATTTCTTTGGAAAGTGAAGCAAACCTAAACTCAGATAAA
GAAAATATAACCACCTCAAATCTCAAGGCGAGTCATTCCCCTCCTTTGAATCTACCCAACAA
CAGCCACGGAATAACAGATTTCTCCAGTAACATCAGCAGAGCATTCTTTGGGCAGTCTAA
AACCCACATCTACCATTTCCACAAGCCCTCCCTTGATCCATAGCTTTGTTTCTAAAGTGCCT
TGGAATGCACCTATAGCAGATGAAGATCTTTTGCCCATCTCAGCACATCCCAATGCTACACC
TGCTCTGTCTTCAGAAACTTCACTTGGTCTTTGGTCAATGACACCGTGAAAACCTCCTGATA
ACAGTTCCATTACAGTTAGCATCCTCTCTTCAGAACCAACTTCTCCATCTGTGACCCCTTG
ATAGTGGAACCAAGTGGATGGCTTACCACAAACAGTGATAGCTTCACTGGGTTTACCCCTTA
TCAAGAAAAACAACCTCTACAGCCTACCTTAAAATTCACCAATAATTCAAAACTCTTTCCAA
ATACGTCAGATCCCCAAAAAGAAAATAGAAATACAGGAATAGTATTCGGGGCCATTTTAGGT
GCTATTCTGGGTGTCTCATTGCTTACTCTTGTGGGCTACTTGTGTGTGAAAAAGGAAAAC
GGATTCATTTTCCCATCGGCGACTTTATGACGACAGAAATGAACCAGTCTGCGATTAGACA
ATGCACCGGAACCTTATGATGTGAGTTTTGGGAATTCTAGCTACTACAATCCAACCTTTGAAT
GATTCAGCCATGCCAGAAAGTGAAGAAAATGCACGTGATGGCATTCCCTATGGATGACATACC
TCCACTTCGTACTTCTGTATTAGAACTAACAGCAAAAAGGCGTTAAACAGCAAGTGTCTCTA
CATCCTAGCCTTTTGACAAATTCATCTTTCAAAAGGTTACACAAAATTACTGTCACGTGGAT
TTTGTCAAGGAGAATCATAAAAGCAGGAGACCAGTAGCAGAAATGTAGACAGGATGTATCAT
CCAAAGGTTTTCTTTCTTACAATTTTTGGCCATCCTGAGGCATTTACTAAGTAGCCTTAATT
TGTATTTTAGTAGTATTTTCTTAGTAGAAAATATTTGTGGAATCAGATAAAACTAAAAGATT
TCACCATTACAGCCCTGCCTCATAACTAAATAATAAAAATTATTCCACCAAAAAATTCTAAA
ACAATGAAGATGACTCTTACTGCTCTGCCTGAAGCCCTAGTACCATAATTCAAGATTGCAT
TTTCTTAAATGAAAATTGAAAGGGTGCTTTTTTAAAGAAAATTTGACTTAAAGCTAAAAGAG
GACATAGCCCAGAGTTTCTGTTATTGGGAAATTGAGGCAATAGAAATGACAGACCTGTATTC
TAGTACGTTATAATTTTCTAGATCAGCACACACATGATCAGCCCACTGAGTTATGAAGCTGA
CAATGACTGCATTCAACGGGGCCATGGCAGGAAAGCTGACCCTACCCAGGAAAGTAATAGCT
TCTTTAAAGTCTTCAAAGGTTTTTGGGAATTTTAACTTGTCTTAATATATCTTAGGCTTCAA
TTATTTGGGTGCCTTAAAAACTCAATGAGAATCATGGT

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58732

><subunit 1 of 1, 334 aa, 1 stop

><MW: 36294, pI: 4.98, NX(S/T): 13

MLALAKILLISTLFYSLLSGSHGKENQDINTTQNI AEVFKTMENKPISLESEANLNSDKENI
TTSNLKASHSPPLNLPNNSHGITDFSSNSSAEHSLGSLKPTSTISTSPPLIHSFVSKVPWNA
PIADEDLLPISAHPNATPALSSNFTWVSLVNDTVKTPDNSSITVSILSSEPTSPSVTPLIVE
PSGWLTNTNSDSFTGFTPYQEKTTLQPTLKFTNNSKLFNPNTSDPQKENRNTGIVFGAILGAIL
GVSLLTLVGYLLCGKRKTDVSFHRRLYDDRNEPVLRLDNAPEPYDVVSFGNSSYYNPTLNDSA
MPESEENARDGIPMDDIPPLRTSV

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 235-262

N-glycosylation site.

amino acids 30-34, 61-65, 79-83, 90-94, 148-152, 155-159,
163-167, 218-222, 225-229, 298-302, 307-311

AACAGGATCTCCTCTTGTCAGTCTGCAGCCCGAGGACGCTGATTCCAGCAGCGCCTTACCGCGC
AGCCCCGAAGATTCACTATGGTGAAAATCGCCTTCAATACCCCTACCGCCGTGCAAAAGGAGG
AGGCGCGGCAAGACGTGGAGGCCCTCCTGAGCCGCACGGTCAGAACTCAGATACTGACCGGC
AAGGAGCTCCGAGTTGCCACCCAGGAAAAAGAGGGCTCCTCTGGGAGATGTATGCTTACTCT
CTTAGGCCCTTTCAATTCATCTTGGCAGGACTTATTGTTGGTGGAGCCTGCATTTACAAGTACT
TCATGCCCCAAGAGCACCATTTACCGTGGAGAGATGTGCTTTTTTGATTCTGAGGATCCTGCA
AATTCCTTTCGTGGAGGAGAGCCTAACTTCCTGCCTGTGACTGAGGAGGCTGACATTCGTGA
GGATGACAACATTGCAATCATTGATGTGCCTGTCCCCAGTTTCTCTGATAGTGACCCTG¹CAG
CAATTATTCATGACTTTGAAAAGGGAATGACTGCTTACCTGGACTTGTTGCTGGGGA²ACTGC
TATCTGATGCCCCCTCAATACTTCTATTGTTATGCCTCCAAAAAATCTGGTAGAGCTCTTTGG
CAAAC³TGGCGAGTGGCAGATATCTGCCTCAAAC⁴TATGTGGTTCGAGAAGACCTAGTTGCTG
TGGAGGAAATTCGTGATGTTAGTAACCTTGGCATCTTTATTTACCAACTTTGCAATAACAGA
AAGTCCTTCCGCCTTTCGTGCGAGAGACCTCTTGCTGGGTTTCAACAAACGTGCCATTGATAA
ATGCTGGAAGATTAGACACTTCCCCAACGAATTTATTGTTGAGACCAAGATCTGTCAAGAG⁵T
AAGAGGCAACAGATAGAGTGTCTTGGTAATAAGAAGTCAGAGATTTACAATATGACTTTAA
CATTAAGGTTTATGGGATACTCAAGATATTTACTCATGCATTTACTCTATTGCTTATGCTTT
AAAAAAAGGAAAAAAAAAAAAAAAAACTACTAACCCTGCAAGCTCTTGTCAAATTTTAGTTTAAT
TGGCATTGCTTGTTTTTTTGA⁶AACTGAAATTACATGAGTTTCATTTTTTCTTTGCATTTATAG
GGTTTAGATTTCTGAAAGCAGCATGAATATATCACCTAACATCCTGACAATAAATTCCATCC
GTTGTTTTTTTTTGTTTGTTTGTTTTTCTTTTCCTTTAAGTAAGCTCTTTATTCATCTTATG
GTGGAGCAATTTTAA⁷AATTTGAAATATTTTAAATTGTTTTTGA⁸ACTTTTTGTGTAA⁹ATATA
TCAGATCTCAACATTGTTGGTTTCTTTTGTTTTTCATTTTGTACA¹⁰ACTTTCTTGAATTTAGA
AATTACATCTTTGCAGTTCTGTTAGGTGCTCTGTAATTAACCTGACTTATATGTGAACAATT
TTCATGAGACAGTCATTTTTTAACTAATGCAGTGATTCTTTCTCACTACTATCTGTATTGTGG
AATGCACAAAATTGTGTAGGTGCTGAATGCTGTAAGGAGTTTAGGTTGTATGAATTCTACAA
CCCTATAATAAATTTTACTCTATACAAAAAAAAAAAAAAAAAAAAA

FIGURE 26

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58828

<subunit 1 of 1, 263 aa, 1 stop

<MW: 29741, pI: 5.74, NX(S/T): 1

MVKIAFNTPTAVQKEEARQDVEALLSRTVRTQILTGKELRVATQEKEGSSGRCMLTLLGLSF
ILAGLIVGGACIYKYFMPKSTIYRGEMCFFDSEDPANSLRGGEPNFLPVTEEADIREDDNIA
IIDVPVPSFSDSDPAAI IHDFEKGMTAYLDLLLGNCYLMPLNTSIVMPPKNLVELFGKLASG
RYLPQTYVVREDLVAVEEIRDVSNLGFIFYQLCNNRKSFRLRRRDLLLGFNKRAIDKCWKIR
HFPNEFIVETKICQE

Type II transmembrane domain:

amino acids 53-75

N-glycosylation site.

amino acids 166-170

Casein kinase II phosphorylation site.

amino acids 35-39, 132-136, 134-138

N-myristoylation site.

amino acids 66-72, 103-109

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 63-74

FIGURE 27

GGAGGAGGGAGGGCGGGCAGGCGCCAGCCCAGAGCAGCCCCGGGCACACGGA

CTTCCAGCCCAGGTGCCCCCACTCTCGCTCCATTTCGGCGGGAGCACCCAGTCTGTACGCC

AAGGAACTGGTCCTGGGGGCACC**ATG**GTTTCGGCGGCAGCCCCAGCCTCCTCATCCTTCTG

TTGCTGCTGCTGGGGTCTGTGCCTGCTACCGACGCCCGCTCTGTGCCCTGAAGGCCACGTT

CCTGGAGGATGTGGCGGGTAGTGGGGAGGCCGAGGGCTCGTCGGCCTCCTCCCCGAGCCTCC

CGCCACCCCTGGACCCCGGCCCTCAGCCCCACATCGATGGGGCCCCAGCCCACAACCCTGGGG

GGCCCATCACCCCCACCAACTTCCTGGATGGGATAGTGGACTTCTTCCGCCAGTACGTGAT

GCTGATTGCTGTGGTGGGCTCCCTGGCCTTTCTGCTGATGTTTCATCGTCTGTGCCGCGGTCA

TCACCCGGCAGAAGCAGAAGGCCTCGGCCTATTACCCATCGTCCTTCCCCAAGAAGAAGTAC

GTGGACCAGAGTGACCGGGCCGGGGGCCCCCGGGCCTTCAGTGAGGTCCCCGACAGAGCCCC

CGACAGCAGGCCCCGAGGAAGCCCTGGATTCTCCCGGCAGCTCCAGGCCGACATCTTGGCCG

CCACCCAGAACCTCAAGTCCCCCACCAGGGCTGCACTGGGCGGTGGGGACGGAGCCAGGATG

GTGGAGGGCAGGGGCGCAGAGGAAGAGGAGAAGGGCAGCCAGGAGGGGGACCAGGAAGTCCA

GGGACATGGGGTCCCAGTGGAGACACCAGAGGCGCAGGAGGAGCCGTGCTCAGGGGTCTTG

AGGGGGCTGTGGTGGCCGGTGAGGGCCAAGGGGAGCTGGAAGGGTCTCTCTTGTTAGCCAG

GAAGCCCAGGGACCAGTGGGTCCCCCGAAAGCCCCTGTGCTTGACAGCAGTGTCCACCCAG

TGTC**TAA**CAGTCTCCCGGGCTGCCAGCCCTGACTGTCGGGGCCCCCAAGTGGTCACCTCCCC

GTGTATGAAAAGGCCTTCAGCCCTGACTGCTTCCTGACACTCCCTCCTTGGCCTCCCTGTGG

TGCCAATCCCAGCATGTGCTGATTCTACAGCAGGCAGAAATGCTGGTCCCCGGTGCCCCGGA

GGAATCTTACCAAGTGCCATCATCCTTCACCTCAGCAGCCCCAAAGGGCTACATCCTACAGC

ACAGCTCCCCTGACAAAGTGAGGGAGGGCACGTGTCCCTGTGACAGCCAGGATAAAACATCC

CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCCCAAATACTTTTTTAAACA

GCTACAGGGTAAAATCCTGCAGCACCCACTCTGGAAAATACTGCTCTTAATTTTCCTGAAGG

TGGCCCCCTGTTTCTAGTTGGTCCAGGATTAGGGATGTGGGGTATAGGGCATTTAAATCCTC

TCAAGCGCTCTCCAAGCACCCCGGCCTGGGGGTGAGTTTCTCATCCCGCTACTGCTGCTGG

GATCAGGTTGAATGAATGGAACCTTTCCTGTCTGGCCTCAAAGCAGCCTAGAAGCTGAGGG

GCTGTGTTTGAGGGGACCTCCACCCTGGGGAAGTCCGAGGGGCTGGGGAAGGGTTTCTGACG

CCCAGCCTGGAGCAGGGGGGGCCCTGGCCACCCCTGTTGCTCACACATTGTCTGGCAGCCTG

TGTCCACAATATTCGTCACTCCTCGACAGGGAGCCTGGGCTCCGTCTGCTTTAGGGAGGCT

CTGGCAGGAGGTCTCTCCCCCATCCCTCCATCTGGGGCTCCCCAACCTCTGCACAGCTCT

CCAGGTGCTGAGATATAATGCACCAGCACAAATAAACCTTTATTCCGGCCTGAAAAAAAAAAAA

AAGA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58852

><subunit 1 of 1, 283 aa, 1 stop

><MW: 29191, pI: 4.52, NX(S/T): 0

MVSAAPSLILLILLLLLLGSPATDARSVPLKATFLEDVAGSGEAGSSASSPSLPPPWTPAL
SPTSMGPQPTTLGGPSPPTNFLDGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQKQKA
SAYYPSSFPPKKKYVDQSDRAGGPRAFSEVPDRAPDSRPEEALDSSRQLQADILAATQNLKSP
TRAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVPVETPEAQEEPCSGVLEGAVVAGE
GQGELEGSLLLAQEAQGPVGPPESPCACSSVHPSV

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 94-118

N-myristoylation site.

amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199,
211-217, 238-244, 242-248

FIGURE 29

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGA
 GGGAGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCG
 GCAAGGAGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAG**ATG**
 AAGTTCCAGGGGGCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGG
 CCCCCTGCAGAGCGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCC
 TGGGAGACGCCCTGAGCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCT
 GGCTCTAAAGTCAGTGAGGGCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAG
 GCAGGTTCCAGGCTTTGGCGCAGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATG
 CTCTGGGAAACACTGGGCACGAGATTGGCAGACAGGCAGAAGATGTCATTCGACACGGAGCA
 GATGCTGTCCGCGGCTCCTGGCAGGGGGTGCCTGGCCACAGTGGTGCCTTGGGAAACTTCTGG
 AGGCCATGGCATCTTTGGCTCTCAAGGTGGCCTTGGAGGCCAGGGCCAGGGCAATCCTGGAG
 GTCTGGGGACTCCGTGGGTCCACGGATACCCCGGAAACTCAGCAGGCAGCTTTGGAATGAAT
 CCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGAGGGCCACCAAACCTTTGGGACCAACAC
 TCAGGGAGCTGTGGCCCAGCCTGGCTATGGTTCAGTGAGAGCCAGCAACCAGAATGAAGGGT
 GCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCAACTCTGGGGGAGGCAGCGGC
 TCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAACAATGGCAGCAGCAGTGG
 TGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAGTGGCGGCAGCAGTG
 GTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGTCCTCCTGGGGA
 TCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGACATAAACC
 CGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGAGCGGGGAATCTGGGATTCAGGGCTTCA
 GAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTGGGA
 GGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGT
 TGGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCT
 GGAAGAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGA
 AGCTCTCGCATCCCG**TGA**CCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACT
 CCCTCCTTAAACACCACCCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTA
 GCTGCCCCACAAA
 AA

FIGURE 30

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212
```

```
><subunit 1 of 1, 440 aa, 1 stop
```

><MW: 42208, pI: 6.36, NX(S/T): 1

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
 AGSKVSEALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
 ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGLGGQGQGNPGGLGTPWVHGYPGNSAGSFGM
 NPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGS
 GSQSGSSSGSGSNGDNNNGSSSGSSSGSSSGSSSGSSSGSSSGSSSGNSGGSRGDSGSESSW
 GSSTGSSSGNHGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRL
 GSGSDNYRGQGSSWGSGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQ
 RSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 31

GACCGGTCCCTCCGGTCTGGATGTGCGGACTCTGCTGCAGCGAGGGCTGCAGGCCCGCCGGGCGGTGCTCACCCG
 TGCCCTGGCTGGTGGAGTTTCTCTCCTTTGCTGACCATGTTGTTCCCTTGCTGGAATATTACCGGGACATCTTCA
 CTCTCCTGCTGCGCCTGCACCGGAGCTTGGTGTGTGTCGAGGAGAGTGAGGGGAAGATGTTTCTGAAACAAGC
 TGCTGCTACTTGCTGTCTGGGCTGGCTTTTCCAGATTCCACAGTCCCTGAGGACTTGTTCTTTCTGGAAGAGG
 GTCCCTCATATGCCTTTGAGGTGGACACAGTAGCCCCAGAGCATGGCTTGACAATGCGCCTGTGGTGGACCAGC
 AGCTGCTCTACACCTGCTGCCCCCTACATCGGAGAGCTCCGGAAACTGCTCGCTTCGTGGGTGTCAGGCAGTAGTG
 GACGGAGTGGGGGCTTCATGAGGAAAATCACCCCCACCACTACCACCAGCCTGGGAGCCCAGCCTTCCCAGACCA
 GCCAGGGGCTGCAGGCACAGCTCGCCCAGGCCCTTTTCCACAACCAGCCGCCCTCCTTGCGCCGACCGTAGAGT
 TCGTGCGCAGAAAGAAATTGGATCAAACCTGTGTCAAACATATCAAGGCTACACTGGTGGCAGATCTGGTGCGCCAGG
 CAGAGTCACTTCTCCAAGAGCAGCTGGTGACACAGGGAGAGGAAGGGGGAGACCCAGCCCAGCTGTTGGAGATCT
 TGTGTTCCAGCTGTGCCCTCACGGGGCCCAGGCATTGGCCCTGGGGCGGGAGTTCTGTCAAAGGAAGAGCCCTG
 GGGCTGTGCGGGCGCTGCTTCCAGAGGAGACCCCGGCAGCCGTTCTGAGCAGTGCAGAGAACATTGCTGTGGGGC
 TTGCAACAGAGAAAGCCTGTGCTTGGCTGTGAGCCAACATCACAGCACTGATCAGGAGGGAGGTGAAAGCAGCAG
 TGAGTCGCACACTTCGAGCCCAGGGTCTGAACCTGCTGCCCGGGGGAGCGGAGGGGCTGCTCCCGCGCCTGAC
 GTGCTCTCCTTGGCCGTGGGGCCACGGGACCTGACGAGGGAGTCTCCCAGAGCATCTGGAACAGCTCCTAGGC
 CAGCTGGGCCAGACGCTGCGGTGCCGCCAGTTCTGTGCCACCTGCTGAGCAGCATCTGGCAAAGTGTCTGTG
 GAGTTAGCTTCCCTCCTCGTTGCAGATCAAATTCCTATCCTAGGGCCCCCGGCACAGTACAGGCTGGAGAGAGGG
 CAGGCTCGAAGGCTTCTGCACATGCTGCTTTCTTGTGGAAGGAAGACTTTCAGGGGCCGTTCCGCTGCAGCTG
 CTGCTGAGCCCAAGAAATGTGGGGCTTCTGGCAGACACAAGGCCAAGGGAGTGGGACTTGCTGCTATTCTTGCTA
 CGGGAGCTGGTGGAGAAGGGTCTGATGGGACGGATGGAGATAGAGGCCTGCCTGGGCAGCCTCCACCAGGCCAG
 TGGCCAGGGGACTTTGCTGAAGAATTAGCAACACTGTCTAATCTGTTTCTAGCCAGCCCCACCTGCCAGAAGCC
 CAGCTAAGAGCCTGTGAGTTGGTGCAGCCAAACCGGGGCAGTGTGCTGGCCAGAGCTAGGGCTGAGAAGTGGCC
 CTGCTTGGGCATTGCACCAGAACCCTGGACCCCCGCTCACGAGGAGGCCCAAGTGCCCAATGCAGACCCCTCAC
 TGGTTGGGGTGTAGCTGGGTCTACAGTCACTTCTGCTCTAAGGGTGTCACTGCCTGGCATCCACCACGCGA
 ATCCTAGAGGAAGGAGAGTTGGCCTGATTTGGGATTATGGCAGAAAAGTCCAGAGATGCCAGTCTCGGAGTAGAA
 GAGGTGGTGTGTTTATCTCTTGGATACTAAATGAAATGAGGTGTGTGGGCTTGTCAACACAGAATTCAAGCCT
 CATTTGCTATCCCAGCATCTCTTAAACCTTTGTAGTCTTGGAAATCATGACAGAGGCAAATGACTCCTGCTTAAC
 TTATGAAGAAAGTTAAACATGAATCTTGGGAGTCTACATTTCTTATCACCAGGAGCTGGACTGCCATCTCCTT
 ATAAATGCCTAACACAGGCCGGGTCTGGTGGCTCATGCCTGTAATCCCAGCACTTTGAGAGGCCTGAGGTGGCG
 GACTGCCTGAGGTGAGGAATTCAAGACCAGCCTGGCCAACATGGCAAACCCCATCTCTACTAAAAATAAAAAA
 TTATTAGCTGGGCATGGTGGTGTGTGCTGTAATCCCAGCTACTCAGGAGGATGAGGCAGGAGACCTGCTTGAAC
 CTGGAGGTGGAGGTGTCAGTGAGCCGAGGTGCGACCACTGCACTCCAGTCTGGGTAACAGAGCGAGACTTTCTAG
 AAAAAGCCTAACAAACAGATAAGGTAGGACTCAACCACTGAAACCTGACTTTCCCCCTGTACCTTCAGCCCCCTG
 TGCAGGTAGTAACCTCTTGAACCTCTCCCTGACCAGGGACCAAGCACAGGGCATTTAGAGCTTTTTAGAAATAA
 CTGGTTTTCTTTAAAAAAGGGCTTTTATTAATAATCTCCCCACACGATGGCTCCTGCAATCTGCCACAGCTC
 TTTTTTTTTTTTTTTTTTAAAAAGGGCTTTTATTAATAATCTCCCCACACGATGGCTCCTGCAATCTGCCACAGCTC
 TGGGGCGTGTCTGTAGGGAAAGGCCCTGTTTTCCCTGAGGCGGGGCTGGGCTTGTCATGGGTCCGCGGAGCTG
 GCCGTGCTTGGCGCCCTGGCGTGTGTCTAGCTGCTTCTTGGCGGCACAGAGCTGCGGGGTCTGGGGGACCGGG
 AGCTAAGAGCAGGCTCTGGTGCAGGGGTGGAGGCCTGTCTTAACCGACACCCTGAGGTGCTCCTGAGATGCTG
 GGTCCACCCTGAGTGGCACGGGGAGCAGCTGTGGCCGGTGTCTCTCYTAGGCCAGTCTGGGGAAACTAAGCTC
 GGGCCCTTCTTTGCAAAGACCGAGGATGGGGTGGGTGTGGGGGACTCATGGGAATGGCCTGAGGAGCTACGTGT
 GAAGAGGGCGCCGGTTTGTGGCTGCAGCGGCCTGGAGCGCCTCTCTCTGAGCCTCAGTTCCCTTTCCGTCTA
 ATGAAGAACATGCCGTCTCGGTGTCTCAGGGCTATTAGGACTTGCCCTCAGGAAGTGGCCTTGGACGAGCGTCAT
 GTTATTTTTACAACCTGTCCTGCGACGTTGGCCTGGGCACGTGGAATGGCCATGTCCCTCTGCTGCGTGGAC
 GTCGCGGTGCGGAGTGCAGCCAGAGGGCGGGGCCAGACGTGCGCCTGGGGGTGAGGGGAGGCGCCCCGGGAGGG
 CCTCACAGGAAGTTGGGCTCCCGCACCAAGGCAGGGCGGGCTCCCGCCGCGCCGCGCCACCACCGTCCAGG
 GGCCGGTAGACAAAGTGGAAGTGCAGCTTGGGCTCGCTGCGCAGCAGGTAGCCCTTGATGCAGTGCGGCAGCGC
 TCGTCCGCCAGCTGGAAGCAGCGCCCTCCACCAGCACGAACAGCCGGTGCCT

FIGURE 32

MCFLNKL L L L L L A V L G W L F Q I P T V P E D L F F L E E G P S Y A F E V D T V A P E H G L D N A P V V D Q Q L L Y T C
C P Y I G E L R K L L A S W V S G S S G R S G G F M R K I T P T T T T S L G A Q P S Q T S Q G L Q A Q L A Q A F F H N Q P P
S L R R T V E F V A E R I G S N C V K H I K A T L V A D L V R Q A E S L L Q E Q L V T Q G E E G G D P A Q L L E I L C S Q L
C P H G A Q A L A L G R E F C Q R K S P G A V R A L L P E E T P A A V L S S A E N I A V G L A T E K A C A W L S A N I T A L
I R R E V K A A V S R T L R A Q G P E P A A R G E R R G C S R A

Signal peptide:

amino acids 1-18

N-glycosylation site.

amino acids 244-248

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 89-93

Casein kinase II phosphorylation site.

amino acids 21-25, 167-171, 223-227

N-myristoylation site.

amino acids 100-106, 172-178, 207-213

Microbodies C-terminal targeting signal.

amino acids 278-282

TCCCTTGACAGGTCTGGTGGCTGGTTCTGGGGTCTACTGAAGGCTGTCTTGATCAGGAAACTG
AAGACTCTCTGCTTTTGCCACAGCAGTTCCTGCAGCTTCCTTGAGGTGTGAACCCACATCCC
TGCCCCCAGGGCCACCTGCAGGACGCCGACACCTACCCCTCAGCAGACGCCGGAGAGAA**ATG**
AGTAGCAACAAAGAGCAGCGGTGAGCAGTGTTCGTGATCCTCTTTGCCCTCATCACCATCCT
CATCCTCTACAGCTCCAACAGTGCCAATGAGGTCTTCCATTACGGCTCCCTGCGGGGCCGTA
GCCGCCGACCTGTCAACCTCAAGAAGTGGAGCATCACTGACGGCTATGTCCCCATTCTCGGC
AACAAAGACACTGCCCTCTCGGTGCCACCAGTGTGTGATTGTGACGAGCTCCAGCCACCTGCT
GGGCACCAAGCTGGGCCCTGAGATCGAGCGGGCTGAGTGTACAATCCGCATGAATGATGCAC
CCACCACTGGCTACTCAGCTGATGTGGGCAACAAGACCACCTACCGCGTCGTGGCCCATTC
AGTGTGTTCCGCGTGCTGAGGAGGCCCCAGGAGTTTGTCAACCGGACCCCTGAAACCGTGTT
CATCTTCTGGGGGGCCCCCGAGCAAGATGCAGAAGCCCCAGGGCAGCCTCGTGCCTGTGATCC
AGCGAGCGGGCCTGGTGTTCCTCCCAACATGGAAGCATATGCCGTCTCTCCCGGCCGCATGCGG
CAATTTGACGACCTCTTCCGGGGTGAGACGGGCAAGGACAGGGAGAAGTCTCATTTCGTGGTT
GAGCACAGGCTGGTTTACCATGGTGATCGCGGTGGAGTTGTGTGACCACGTGCATGTCTATG
GCATGGTCCCCCCCCAACTACTGCAGCCAGCGGCCCCGCCTCCAGCGCATGCCCTACCACTAC
TACGAGCCCCAAGGGGGCCGGACGAATGTGTACCTACATCCAGAATGAGCACAGTCGCAAGGG
CAACCACCACCGCTTCATCACCGAGAAAAGGGTCTTCTCATCGTGGGCCCAGCTGTATGGCA
TCACCTTCTCCACCCCTCCTGGACC**TAG**GCCACCCAGCCTGTGGGACCTCAGGAGGGTCAG
AGGAGAAGCAGCCTCCGCCCAGCCGCTAGGCCAGGGACCATCTTCTGGCCAATCAAGGCTTG
CTGGAGTGTCTCCAGCCAATCAGGGCCTTGAGGAGGATGTATCCTCCAGCCAATCAGGGCC
TGGGGAATCTGTTGGCGAATCAGGGATTTGGGAGTCTATGTGGTTAATCAGGGGTGTCTTTC
TTGTGCAGTCAGGGTCTGCGCACAGTCAATCAGGGTAGAGGGGGTATTTCTGAGTCAATCTG
AGGCTAAGGACATGTCCCTTTCCCATGAGGCCTTG GTTCAGAGCCCCAGGAATGGACCCCCCA
ATCACTCCCCACTCTGCTGGGATAATGGGGTCTGTCCCAAGGAGCTGGGAACTTG GTGTTG
CCCCCTCAATTTCCAGCACACCAGTTCAGTGCGGAGTCTGAGACCCTTTCCACCTCCCCTACAA
GGCCAGAGAAATTTGTGGGGTGTGTGGAGGTGTGGGGGCGGTGGGGAGGTCCCAGAGGTGGGA
GGCTGGCATCCAGGTCTTGGCTCTGCCCTGAGACCTTGGACAAACCCTTCCCCCTCTCTGGG
CACCCTTCTGCCACACCAGTTTCCAGTGCGGAGTCTGAGACCCTTTCCACCTCCCCTACAA
GTGCCCTCGGGTCTGTCTCTCCCGTCTGGACCCTCCCAGCCACTATCCCTTGCTGGAAGGCT
CAGCTCTTTGGGGGGTCTGGGGTGACCTCCCCACCTCCTGGAAAACTTTAGGGGTATTTTTC
GCAAACCTCCTTCAGGGTTGGGGGACTCTGAAGGAAACGGGACAAAACCTTAAGCTGTTTTCT
TAGCCCCCTCAGCCAGCTGCCATTAGCTTGGCTCTTAAAGGGCCAGGCCTCCTTTTCTGCCCT
CTAGCAGGGGAGGTTTTTCCAACCTGTTGGAGGCGCCTTTTGGGGCTGCCCTTTGTCTGGAGTCA
CTGGGGGCTTCCGAGGGTCTCCCTCGACCCTCTGTCTGTGGGATGGCTGTCTGGGAGCTGT
ATCACCTGGGTTCTGTCCCCTGGCTCTGTATCAGGCACCTTTATTAAAGCTGGGCCTCAGTGG
GGTGTGTTTTGTCTCCTGCTCTTCTGGAGCCTGGAAGGAAAGGGCTTCAGGAGGAGGCTGTGA
GGCTGGAGGGACCAGATGGAGGAGGCCAGCAGCTAGCCATTGCACACTGGGGTGATGGGTGG
GGGCGGTGACTGCCCCAGACTTG GTTTTTGTAATGATTTGTACAGGAATAAACACACCTACGC
TCCGGAAAAA

FIGURE 34

MSSNKEQRSAVFVILFALITILILYSSNSANEVFHYGSLRGRSRRPVNLKKWSITDGYVPIL
GNKTLPSRCHQCIVIVSSSSHLLGTKLGPEIERAECTIRMNDAPTTGYSADVGNKTTYRVVAH
SSVFRVLRRPQEFVNRTPETVFIFWGPSPKMQKPQGSLVRVIQRAGLVFPNMEAYAVSPGRM
RQFDDLFRGETGKDREKSHSWLSTGWFTMVI AVELCDHVHVYGMVPPNYCSQRPRLQRMPYH
YYEPKGPDECVTYIQNEHSRKGNNHHRFITEKRVFSSWAQLYGITFSHPSWT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 9-31 (type II)

N-glycosylation site.

amino acids 64-68, 115-119

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 50-54

Casein kinase II phosphorylation site.

amino acids 3-7, 29-33, 53-57, 197-201

Tyrosine kinase phosphorylation site.

amino acids 253-262

N-myristoylation site.

amino acids 37-43, 114-120, 290-294

GTTTCTCATAGTTGGCGTCTTCTAAAGGAAAAACACTAAAATGAGGAACCTCAGCGGACCGGGAGCGACGCAGCTT
GAGGGAAGCATCCCTAGCTGTTGGCGCAGAGGGGCGAGGCTGAAGCCGAGTGGCCCGAGGTGTCTGAGGGGCTGG
GGCAAAGGTGAAGAGTTCAGAAACAAGCTTCCTGGAAACCCATGACCCATGAAGTCTTGTCGCACATTTATACCGT
CTGAGGGTAGCAGTTCGAACTAGAAGAGTGGAGTGTGCCAGGACGGCAGTATCTCTTTGTGTGACCTGGC
GGCCTATGGGACGTTGGCTTCAGACCTTTGTGATACACCA**ATGCT**CGCTGGGACGATGACGGCGTGGAGGGAATG
AGGCCTGAGGTACACTGGCTTGCTCCTCCTAGCCACAGCAGGCTGCTTTGCTGACTTGAACGAGGTCCCTCAG
GTCACCGCTCCAGCGTGCCTCCACCGTCCGAAAGCCCGGAGGCACCTGTGATCTTTGGGCTGCGCTGGTGGAACTCCA
AGGATGAATGTAACCTGGCGCTGAATGGAAGGAGCTGAATGGCTCGGATGATGCTCTGGGTGCTCCTCATACC
CACGGGACCTCGTCATCATCTGCCCTTAACAACCACTGTGGGACGTTACGAGTGTGTGGCCCGATGCCCTGCG
GGGCTGTGGCCAGCGTGCCAGCCACTGTGACACTAGCCAATCTCCAGGACTTCAAGTTAGATGTGCAGCACGT
ATTGAAGTGGATGAGGGAACACAGCAGTCATTGCCTGCCACCTGCCTGAGAGCCACCCCAAAGCCAGGTCCGG
TACAGCGTCAACAAGAGTGGCTGGAGGCTCCAGAGGTAACCTACCTGATCATGCTCCCTCAGGGAACCTCCAGATT
GTGAATGCCAGCCAGGAGGACGAGGGCATGTACAAGTGTGACGCTACAACCCAGTGACCCAGGAAGTGAACAC
TCCGGCTCCAGCGACAGGCTACGTGTGCGCCGCTCCACCGCTGAGGCTGCCCGCATCATCTACCCCCAGAGGCC
CAAACCATCATCGTCACCAAAGGCCAGAGTCTCATTCTGGAGTGTGTGGCCAGTGGAAATCCCACCCCCACGGGTC
ACCTGGGCGAAGGATGGTCCAGTGTACCCGGCTACACAAGACGCGCTTCCTGCTGAGCAACCTCCTCATCGAC
ACCACGCGAGGAGGACTCAGGCACCTACCGCTGCATGGCCGACAATGGGTTGGGACGCCGGGACGGGTCAGGCT
ATCCTCTACAATGTCCAGGTGTTTGAACCCCCCTGAGGTACCATGGAGCTATCCAGCTGGTTCATCCCCGGGGC
CAGAGTGGCAAGCTTACCTGTGAGGTGCGTGGGAACCCCCCGCCCTCGCTGCTGTGGTGTAGGAATGCTGTGCC
CTCATCTCCAGCCAGCGCTCCGGCTCTCCCGCAGGGCCCTGCGCTGCTCAGCATGGGGCTGTAGGACGAAGGC
GTCTACCAAGTCATGGCCGAGAACGAGGTTGGGAGCGCCATGCCGTAGTCCAGCTGCGGACCTCCAGGCCAAGC
ATAACCCCAAGGCTATGGCAGGATGCTGAGCTGGCTACTGGCACACCTCCTGTATACCCCTCAAACCTCGGCAAC
CCTGAGCAGATGCTGAGGGGGCAACCGCGCTCCCCAGACCCCCAACGTCAGTGGGGCTGCTTCCCCGAAGTGT
CCAGGAGAGAAGGGCAGGGGCTCCCGCGAGGCTCCCATCTACCTCAGCTCGCCCCGCACCTCCAAGACAGAC
TCATATGAAGTGGTGTGGCGGCTCGGCATGAGGGCAGTGGCGGGCGCCAATCCTCTACTATGTGGTGAACAC
CGCAAGCAGGTCAAAATTCCTCTGACGATTGGACCATCTCTGGCATTCAGGCCAACAGCAGCCGCTGACCCCTC
ACCGACTTGACCCCGGGAGCTTGATGAAGTGGAGTGGCAGCTTACAACCTGTGCGGGAGAGGGCCAGACAGCC
ATGGTACCTTCCGAACCTGGACGGCGGCCAAACCCGAGATCATGGCCAGCAAGAGCAGAGATCCAGAGAGAG
GACCTGGAGCCAGTCCCCAGAGCAGCAGCCAGCCAGACCAGCGCGCTCTCCCCCAAGAGCTCCCCGACAGG
CCCACCATCTCCACGGCCTCCGAGACCTCAGTGTACGTGACCTGGATTCCCCGTGGGAATGGTGGGTTCCCAATC
CAGTCTTCCGTGTGGAGTACAAGAAGCTAAAGAAAGTGGGAGACTGGATTCTGGCCACCAGCGCCATCCCCCA
TCGCGGCTGTCCGTGGAGATCACGGGCTTAGAGAAAGGCACCTCCTACAAGTTTCGAGTCCGGGCTCTGAACATG
CTGGGGAGAGCGAGCCAGCGCCCCCTCTCGGCCCTACGTGGTGTCTGGGCTACAGCGGTCGCGTGTACGAGAGG
CCCGTGGCAGGTCTTTATATCACCTTTCACGGATGCGGTCAATGAGACCACCATCATGCTCAAGTGGATGTACATC
CCAGCAAGTAACAACAACACCCCAATCCATGGCTTTTATATCTATTATCGACCCACAGACAGTGACAATGATAGT
GACTACAAGAAGGATATGGTGGAAAGGGCAAGTACTGGCACTCCATCAGCCACCTGCAGCCAGGACCTCCTAC
GACATTAAGATGCAGTGCTTCAATGAAGGAGGGGAGAGCGAGTTCAGCAACCTGATGATCTGTGAGACCAAGCT
CGGAAGTCTTCTGGCCAGCCTGGTGCAGTGCCACCCCCAACTCTGGCCCCACCACAGCCGCCCTTCTCTGAAACC
ATAGAGCGGCGCGGTGGGCATGGGGCCATGGTGGCTCGCTCCAGCGACCTGCCCTATCTGATTGTGCGGGTCTGT
CTGGGCTCCATCGTTCTCATCATCTCGTACCTTCTATCCCTTCTGCTTGTGGAGGGCTGGTCTAAGCAAAAAACAT
ACAACAGACCTGGGTTTTCTCGAAGTGCCTTCCACCTCTCTGCCGTATACATGTGTGCCATTGGGAGGACTC
CCAGGCCACCAGGCCAGTGGACAGCCCTACCTCAGTGGCATCAGTGGACGGGCTGTGCTAATGGGATCCACATG
AATAGGGGCTGCCCTCGGCTGCACTGGGTACCCGGGCATGAAGCCCCAGCAGCAGTCCCGAGCGAGCTTCAG
CAGCAGAGTGACACCAGCAGCTGCTGAGGCAGACCCATCTTGCAATGGATATGACCCCCAAAGTCAACGATC
ACGAGGGTCCCAAGTCTAGCCGAGCAGAGGGCTCTTCTTATACACACTGCCGACGACTCCACTCACCAGCTG
CTGCAGCCCCATCAGACTGCTGCCAACGCCAGGAGCAGCCTGCTGCTGTGGGCCAGTCAGGGGTGAGGAGAGCC
CCCGACAGTCTGTCTGGAAGCAGTGTGGGACCTCCATTTCACTCAGGCGCCCCATGCTGCTTGGGCTTGTG
CCAGTTGAAGAGGTGGACAGTCTGACTCTGCAAGTGAAGTGGAGGAGCTGGTGTCCCCAGCACCCCGTAGGG
GCTACGTAGGACAGGAACCTGGAATGCAGCTCTCCCGGGGGCAGTGGTGCCTGTGCTTTTGAACACCACTC
CTCACAATTTAGGCAGAAAGTGTATATCCAGAAAGACTATATATTGTTTTTTTTTAAAAAAGAAAGAAAAA
AGAGACAGAGAAAAATGGTATTTATTTTCTATTATAGCCATATTTATATATTTATGTAAGTGTAAATAAATGTA
TATGTTTTATAATTCTGGAGAGACATAAGGAGTCTACCCGTTGAGGTTGGAGAGGCAATAAAGAAAGCTGCCA
CCTAACAGGAGTCACCCAGGAAGCACCGCCACAGGCTGGCGCGGAGCAGACTCCTAACCTGGGGCTCTGCAGTG
GCAGGCGAGGCTGCAGGAGGCCACAGATAAGCTGGCAAGAGGAAGGATCCAGGCACATGGTTCATCACGAGCA
TGAAGGGAACAGCAAGGGGACGGTATCACAGCTGGAGACACCCACAGATGGCTGGATCCGGTGCTACGGGAA
ACATTTCTCAAGATGCCATGAGAACAGCCCAAGATGTGTACGACACTATGACATTAATAAAACCTTCCAGAT
CAATAATCCGTGGCAACATATCTCTGTAAAAACAACACTGTAACCTCTAAATAAATGTTTAGTCTTCCCTGTAAAA

FIGURE 36

MLRGTMTAWRGMRPEVTLACLLLATAGCFADLNEVPQVTVQPASTVQKPGGTIVILGCVVEPP
 RMNVTWRLNGKELNGSDDALGVLITHGTLVITALNNHTVGRYQCVARMPAGAVASVPATVTL
 ANLQDFKLDVQHVIEWDEGNTAVIACHLPESHKPAQVRYSVKQEWLEASRGNYLIMPSTGNLQ
 IVNASQEDEGMYKCAAYNPVTQEVKTSGSSDRLRVRRSTAEAAARI IYPPEAQTI IIVTKGQSL
 ILECVASGI PPRVTWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCMADNGVGQPGA
 AVILYNVQVFEPPEVTMELSOLVIPWGQSAKLTCEVRGNPPPSVLWLRNAVPLISSQRLRLS
 RRALRVLSMGPEDEGVYQCAENEVGSAAHAVVQLRTSRPSITPRLWQDAELATGTPPVSPSK
 LGNPEQMLRGQPALPRPPTSVGPASPKCPGEGKGQAPAEAPI ILSPTSKTDSYELVWRPR
 HEGSGRAPILYYVVKHRKQVTNSSDDWTISGIPANQHRLTLTRLDPGSLYEVEMAAYNCAGE
 GQTAMVTFRTGRRPKPEIMASKEQQIQRDDPGASQSSSQPDHGRLSPPEAPDRPTISTASE
 TSVYVTWIPRGNGGFPIQSFRVEYKKLKKVGDWILATSAIPPSRLSVEITGLEKGTSYKFRV
 RALNMLGESEPSAPSRPYVVS GYSGRVYERPVAGPYITFTDAVNETTIMLKWMIIPASNNNT
 PIHGFYIYYRPTDSDNDSYKKDMVEGDKYWHSISHLQPETS YDIKMQCFNEGGESEFSNVM
 ICETKARKSSGQPGRLPPPTLAPPQPPLPETIERPVGTGAMVARSSDLPYLIVGVVLGSIVL
 IIVTFIPFCLWRAWKQKHTTDLGFPRSALPPSCPYTMVPLGGLPGHQASGQPYLSGISGRA
 CANGIHMNRGCPSAAVGYPGMKPQQHCPGELQQQSDTSSLLRQTHLGNGYDPQSHQITRGPK
 SSPDEGSFLYTLPDDSTHQLLQPHHDCCQRQEQAAGVQSGVRRAPDSPVLEAVWDPPFHSG
 PPCCCLGLVPVEEVDSPDSCQVSGGDWCPQHVPV GAYVGQEPGMQLSPGPLVRVSFETPPLTI

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 16-30 (type II), 854-879

FIGURE 37

CGGGAGGCTGGGTCTGTCATGATCCGGACCCCATTTGTCGGCCTCTGCCCATCGCCTGCTCCTC
CCAGGCTCCCGCGGGCCGACCCCGCGCAAC**ATG**CAGCCCACGGGCCGCGAGGGTTCCCGCGC
GCTCAGCCGGCGGTATCTGCGGCGTCTGCTGCTCCTGCTACTGCTGCTGCTGCTGCGGCAGC
CCGTAACCCGCGCGGAGACCACGCCGGGGCGCCCCCAGAGCCCTCTCCACGCTGGGCTCCCC
AGCCTCTTCACCACGCCGGGTGTCCCCAGCGCCCTCACTACCCCAGGCCTCACTACGCCAGG
CACCCCCAAAACCTTGACCTTCGGGGTTCGCGCGCAGGCCCTGATGCGGAGTTTCCCACTCG
TGGACGGCCACAATGACCTGCCCCAGGTCCTGAGACAGCGTTACAAGAATGTGCTTCAGGAT
GTTAACCTGCGAAATTTAGCCATGGTCAGACCAGCCTGGACAGGCTTAGAGACGGCCTCGT
GGGTGCCCAGTTCTGGTCAGCCTCCGTCTCATGCCAGTCCCAGGACCAGACTGCCGTGCGCC
TCGCCCTGGAGCAGATTGACCTCATTCACCGCATGTGTGCCCTCTACTCTGAACTCGAGCTT
GTGACCTCAGCTGAAGGTCTGAACAGCTCTCAAAGCTGGCCTGCCTCATTGGCGTGNAGGG
TGGTCACTCACTGGACAGCAGCCTCTCTGTGCTGCGCAGTTTCTATGTGCTGGGGGTGCGCT
ACCTGACACTTACCTTCACCTGCAGTACACCATGGGCAGAGAGTTCCACCAAGTTCAGACAC
CACATGTACACCAACGTCAGCGGATTGACAAGCTTTGGTGAGAAAGTAGTAGAGGAGTTGAA
CCGCCTGGGCATGATGATAGATTTGTCCTATGCATCGGACACCTTGATAAGAAGGGTCCTGG
AAGTGTCTCAGGCTCCTGTGATCTTCTCCCACTCAGCTGCCAGAGCTGTGTGTGACAATTTG
TTGAATGTTCCCGATGATATCCTGCAGCTTCTGAAGAACGGTGGCATCGTGATGGTGACACT
GTCCATGGGGGTGCTGCAGTGCAACCTGCTTGCTAACGTGTCCACTGTGGCAGATCACTTTG
ACCACATCAGGGCAGTCATTGGATCTGAGTTCATCGGGATTGGTGGAAATTATGACGGGACT
GGCCGGTTCCCTCAGGGGCTGGAGGATGTGTCCACATACCCAGTCCTGATAGAGGAGTTGCT
GAGTCGTASCTGGAGCGAGGAAGAGCTTCAAGGTGTCCTTCGTGGAAACCTGCTGCGGGTCT
TCAGACAAGTGGAAAAGGTGAGAGAGGAGAGCAGGGCGCAGAGCCCCGTGGAGGCTGAGTTT
CCATATGGGCAACTGAGCACATCCTGCCACTCCCACCTCGTGCCTCAGAATGGACACCAGGC
TACTCATCTGGAGGTGACCAAGCAGCCAACCAATCGGGTCCCCTGGAGGTCCTCAAATGCCT
CCCCATACCTTGTTCCAGGCCTTGTTGGCTGCTGCCACCATCCCAACCTTCACCCAGTGGCTC
TGCT**TGA**CACAGTCCGGTCCCCGCGAGAGGTCACTGTGGCAAAGCCTCACAAAGCCCCCTCTCCT
AGTTCATTACACAAGCATATGCTGAGAATAAACATGTTACACATGGAAAA

FIGURE 38

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59817

><subunit 1 of 1, 487 aa, 1 stop, 2 unknown

><MW: 53569.32, pI: 7.68, NX(S/T): 5

MQPTGREGSRALSRRYLRRLLLLLLLLLLLLRQPVTRAETTPGAPRALSTLGSPSLFTTPGVPS
 ALTTPGLTTPGTPKTLDLRGRAQALMRSPFLVDGHNDLPQVLRQRYKNVLQDVNLRNFSHGQ
 TSLDRLRDGLVGAQFWSASVSCQSQDQTAVRLALEQIDLIHRMCASYSELELVTSAEGLNSS
 QKLACLIGVXGGHSLDSSLSVLRSFYVLGVRYLTLTFTCSTPWAESSTKFRHHMYTNVSGLT
 SFGEKVVEELNRLGMMIDLSYASDTLIRRVLEVSQAPVIFSHSAARAVCDNLLNVPDDILQL
 LKNGGIVMVTLSMGVLQCNNLANVSTVADHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLEDV
 STYPVLIEELLSRXWSEELQGVLRGNLLRVFRQVEKVREESRAQSPVEAEFPYGQLSTSCH
 SHLVPQNGHQATHLEVTKQPTNRVPWRSSNASPYLVPGLVAAATIPTFTQWLC

Important features of the protein:

Signal peptide:

amino acids 1-36

Transmembrane domain:

amino acids 313-331

N-glycosylation sites.

amino acids 119-122, 184-187, 243-246 and 333-336

N-myristoylation sites.

amino acids 41-46, 59-64, 73-78, 133-138, 182-187, 194-199, 324-329, 354-359, 357-362, 394-399, 427-432 and 472-477.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 136-146

Downloaded from www.nrjonline.org by 129.11.24.101 on 01/01/18

FIGURE 39

TGCTAGGCTCTGTCCCACAATGCACCCGAGAGCAGGAGCTGAAAGCCTCTAACACCCACAGA
TCCCTCTATGACTGCAATGTGAGGTGTCCGGCTTTGCTGGCCCAGCAAGCCTGATAAGC**ATG**
AAGCTCTTATCTTTGGTGGCTGTGGTCGGGTGTTTGCTGGTGCCCCAGCTGAAGCCAACAA
GAGTTCTGAAGATATCCGGTGCAAATGCATCTGTCCACCTTATAGAAACATCAGTGGGCACA
TTTACAACCAGAATGTATCCCAGAAGGACTGCAACTGCCTGCACGTGGTGGAGCCCATGCCA
GTGCCTGGCCATGACGTGGAGGCCTACTGCCTGCTGTGCGAGTGCAGGTACGAGGAGCGCAG
CACCACCACCATCAAGGTCATCATTGTCATCTACCTGTCCGTGGTGGGTGCCCTGTTGCTCT
ACATGGCCTTCCTGATGCTGGTGGACCCTCTGATCCGAAAGCCGGATGCATACACTGAGCAA
CTGCACAAATGAGGAGGAGAATGAGGATGCTCGCTCTATGGCAGCAGCTGCTGCATCCCTCGG
GGGACCCCGAGCAAACACAGTCCTGGAGCGTGTGGAAGGTGCCCAGCAGCGGTGGAAGCTGC
AGGTGCAGGAGCAGCGGAAGACAGTCTTCGATCGGCACAAGATGCTCAGC**TAGAT**GGGCTGG
TGTGGTTGGGTCAAGGCCCAACACCATGGCTGCCAGCTTCCAGGCTGGACAAAGCAGGGGG
CTACTTCTCCCTTCCCTCGGTTCCAGTCTTCCCTTTAAAGCCTGTGGCATTTTTCTCCTT
CTCCCTAACTTTAGAAATGTTGTACTTGGCTATTTTGATTAGGGAAGAGGGATGTGGTCTCT
GATCTCTGTTGTCTTCTTGGGTCTTTGGGGTTGAAGGGAGGGGGGAAGGCAGGCCAGAAGGGA
ATGGAGACATTCGAGGCGGCCTCAGGAGTGGATGCGATCTGTCTCTCCTGGCTCCACTCTTG
CCGCCTTCCAGCTCTGAGTCTTGGGAATGTTGTTACCTTGGGAAGATAAAGCTGGGTCTTCA
GGAACTCAGTGTCTGGGAGGAAAGCATGGCCCAGCATTACAGCATGTGTTCCCTTTCTGCAGTG
GTTCTTATCACCACTCCCTCCCAGCCCCGGCGCCTCAGCCCCAGCCCCAGCTCCAGCCCTG
AGGACAGCTCTGATGGGAGAGCTGGGCCCCCTGAGCCCCTGGGTCTTCAGGGTGCAGTGG
AGCTGGTGTTCGCTGTCCCCTGTGCACCTTCTCGCACTGGGGCATGGAGTGCCCATGCATACT
CTGCTGCCGGTCCCCTCACCTGCACCTGAGGGGTCTGGGCAGTCCCTCCTCTCCCCAGTGTC
CACAGTCACTGAGCCAGACGGTCGGTTGGAACATGAGACTCGAGGCTGAGCGTGGATCTGAA
CACCACAGCCCCTGTACTTGGGTGTCCTCTTGTCCCTGAACTTCGTTGTACCAGTGCATGGA
GAGAAAATTTTGTCTCTTGTCTTAGAGTTGTGTGTAAATCAAGGAAGCCATCATTAATTG
TTTTATTTCTCTCA

FIGURE 40

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60278

<subunit 1 of 1, 183 aa, 1 stop

<MW: 20574, pI: 6.60, NX(S/T): 3

MKLLSLVAVVGCLLVPPAEANKSSDIRCKCICPPYRNISGHIYNQNVSQKDCNCLHVVEPM
PVPGHDVEAYCLLCECRYEERSTTTIKVIIIVYLSVVGALLLYMAFLMLVDPLIRKPDAYTE
QLHNEEENEDARSMAAAAASLGGPRANTVLERVEGAQQRWKLQVQEQRKTVFDRHKMLS

Important features:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 90-112

N-glycosylation sites.

amino acids 21-24, 38-41 and 47-50

ACGCGGCTCTCGCTTGGGTTCCGCTAATTTCTGTCTGAGGCGTGAGACTGAGTTTCATAGGGTCCCTGGGTCCCCGA
 ACCAGGAAGGGTTGAGGGAACACAATCTGCAAGCCCCCGGACCCAAGTGAGGGGGCCCCGTGTTGGGGTCCTCCC
 TCCCTTTGCATTCCCACCCCTCCGGGCTTTGCGTCTTCTGGGGACCCCTCGCCGGGAGATGCGCCGCTTGATG
 CGGAGCAAGGATTCGTCTGCTGCCCTCTCTACTGGCCGCGGTGCTGATGGTGGAGAGCTCACAGATCGGCAGT
 TCGCGGGGCAAACTCAACTCCATCAAGTCTCTCTGGCGGGGAGACGCTGGTCAGGCCCAATCGATCTGCG
 GGCATGTACCAAGGACTGGCATTCTGGCGGCAGTAAGAAAGGCAAAAAACCTGGGGCAGGCCTACCCTTGTAGCAGT
 GATAAGGAGTGTGAAGTTGGGAGGTATTGCCACAGTCCCCACCAAGGATCATCGGCCCTGCATGGTGTGTCGGAGT
 AAAAAGAAGCGCTGCCACCGAGATGGCATGTGCTGCCCCAGTACCCGCTGCAATAATGGCATCTGTATCCAGTT
 ACTGAAAGCATCTTAACCCCTCACATCCCGGCTCTGGATGGTACTCGGCACAGAGATCGAAACCACGGTCATTAC
 FCAAACCATGACTTGGGATGGCAGAATCTAGGAAGACCACACATAAGATGTCACATATAAAAGGGCATGAAGGA
 GACCCCTGCCATACGATCATGAGCTGCATTGAAGGGTTTTGCTGTGCTCGTCATTTCTGGACCAAAATCTGCAA
 CCAGTGTCTCCATCAGGGGGAAGTCTGTACCAAACCAAGCAAGGGTTCTCATGGGCTGGAAATTTTCCAGCGT
 TCGGACTGTGCGAAGGGCCTGTCTTGCAAAGTATGGAAAGATGCCACCTACTCCTCCAAAGCCAGACTCCATGTG
 TGTGAGAAAATTTGATCACCATTGAGGAACATCATCAATTGCAGACTGTGAAGTTGTGTATTTAATGCATTATAG
 CATGGTGGAAAATAAGGTTTCAGATGCAGAAGAATGGCTAAAATAAGAAACGTGATAAGAATATAGATGATCACA
 AAAGGGAGAAAAGAAAACATGAACCTGAATAGATTAGAATGGGTGACAAATGCAGTGCAGCCAGTGTTCCTATTATG
 CAACCTGTCTATGTAATAATGTACACATTTGTGGAAAAATGCTATTATTAAGAGAACAAAGCACACAGTGGAAAT
 ACTGATGAGTAGCATGTGACTTTCCAAAGATTTAGGTTGTGCTGGAGGAGAGGTTTCTCTCAGATTGCTGATTGC
 TTATACAAATAACCTACATGCCAGATTTCTATTCAACGTTAGAGTTTAAACAAAATACCTCTAGATAAATCTGTTA
 TACAATAGGTTCTAAAAATAAAAATTTGCTAAACAAGAAATGAAAACATGGAGCATTGTTAATTTACAACAGAAAAAT
 TACCTTTTGATTTGTAACACTACTTCTGCTGTTCAATCAAGAGTCTTGGTAGATAAGAAAAAAATCAGTCAATAT
 TTCCAAATAATTGCAAAATAATGGCCAGTTGTTTAGGAAGGCCTTTAGGAAGACAAATAAATAACAAACAAACAG
 CCACAAATACTTTTTTTTCAAAAATTTTAGTTTTACCTGTAATTAATAAGAACTGATACAAGACAAAAACAGTTCC
 TTCAGATTCTACGGAATGACAGTATATCTCTTTATCCTATTGATTCCTGCTCTGAATGCATTATATTTTCCA
 AACTATATCCCTATAAATTTGTGACTAGTAAATCTTTACACAGAGCAGAATTTTCACAGATGGCAAAAAATTTAA
 GATGTCCAATATATGTGGGAAAAAGAGCTAACAGAGAGATCATTATTTCTTAAAGATTGGCCATAACCTATATTTT
 GATAGAATTAGATTGGTAAATACATGTATTTATACATACTCTGTGGTAATAGAGACTTAAGCTGGATCTGTACTG
 CACTGGAGTAAGCAAGAAAAATTTGGGAAAACTTTTTCGTTTTGTTTCAGGTTTTGGCAACACATAGATCATATGTCTG
 AGGCACAAGTTGGCTGTTCTATCTTTGAAACCAGGGGATGCACAGTCTAAATGAATATCTGCATGGGATTTGCTAT
 CATAATATTTACTATGCAGATGAATTCAGTGTGAGGTCCTGTCCGTACTACTCTCAATTTATTTATTTATAG
 TGCTGAGATCTCTCAAATAATCTCAATTTTCAGGAGGTTTACAAAAATGTACTCCTGAAGTAGACAGACTAGTGAGG
 TTTTCATTGCCCTCTATAAGCTTCTGACTAGCCAATGGCATCATCCAATTTTCTTCCCAAACCTCTGCAGCATCTG
 CTTTATTGCCAAAGGGCTAGTTTCGGTTTTCTGCAGCCATTGCGGTTAAAAAATATAAGTAGGATAACTTGTAAA
 ACCTGCATATTGCTAATCTATAGACACCACAGTTTCTAAATTTCTTTGAAACCACTTTACTACTTTTTTTAAACTT
 AACTCAGTTCTAAATACTTTTGTCTGGAGCACAAAAACAATAAAAGGTTATCTTATAGTCGTGACTTTAAACTTTTG
 TAGACCACAATCTACTTTTGTAGTTTCTTTTACTTAAATCCCATCTGCAGTCTCAAATTTAAGTCTCCAGTAG
 AGATTGAGTTTGAGCCTGTATATCTATTAATAAATTTCAACTTCCCACATATATTTACTAAGATGATTAAAGACTTA
 CATTTTCTGCACAGGTCTGCAAAAAACAAAAATTTATAAACTAGTCCATCCAAGAACCAAGTTTGTATAAACAGGT
 TGCTATAAGCTTGTGAAATGAAATGGAACATTTCAATCAAACATTTCTATATAACAATTATTATATTTACAAT
 TTGGTTCTGCAATATTTTCTTATGTCCACCCTTTTAAAAAATTATTATTTGAAGTAATTTATTTACAGGAAATG
 TTAATGAGATGTATTTTCTTATAGAGATATTTCTTACAGAAAGCTTTGTAGCAGAATATATTTGCAGCTATTGAC
 TTTGTAATTTAGAAAAATGTATAAAGATAAAATCTATTAAATTTTCTCCTCTAAAAACTGAAAAA
 AAAAAA

FIGURE 42

MAALMRSKDSSCCLLLLA AVL MV E S S Q I G S S R A K L N S I K S S L G G E T P G Q A A N R S A G M Y Q G L A
F G G S K K G K N L G Q A Y P C S S D K E C E V G R Y C H S P H Q G S S A C M V C R R K K K R C H R D G M C C P S T R C N N
G I C I P V T E S I L T P H I P A L D G T R H R D R N H G H Y S N H D L G W Q N L G R P H T K M S H I K G H E G D P C L R S
S D C I E G F C C A R H F W T K I C K P V L H Q G E V C T K Q R K K G S H G L E I F Q R C D C A K G L S C K V W K D A T Y S
S K A R L H V C Q K I

Signal peptide:

amino acids 1-25

FIGURE 43

TCTCAATCTAGCTGACCTCGTGATCCGCGCTGACCTTGTAATCCACCTACCTTGGCCTCCCAAA
GTGTTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCAACATCACGTTTTTAAAAATTGATT
TCTTCAAATTCATGGCAAATATTTCCCTTCCCTTTAACTTCTTATGTCAGAATGAGGAAGGA
TAGCTGCATTTATTTAGTCAGTTTTTCATTGCATAGTAATATTTTCATGTAGTATTTTCTAAG
TTATATTTTAGTAATTCATATGTTTTAGATTATAGGTTTTAACATACTTGTGAAAATACTTG
ATGTGTTTTTAAAGCCTTGGGCAGAAATTCTGTATTGTTGAGGATTTGTTCTTTTATCCCCCT
TTTAAAGTCATCCGTCCTTGGCTCAGGATTTGGAGAGCTTGCACCACCAAAAATGGCAAACA
TCACCAGCTCCCAGATTTTGGACCAGTTGAAAGCTCCGAGTTTGGGCCAGTTTACCACCACC
CCAAGTACACAGCAGAATAGTACAAGTCACCCTACAAC TACTACTTCTTGGGACCTCAAGCC
CCCAACATCCCAGTCCTCAGTCCTCAGTCATCTTGACTTCAAATCTCAACCTGAGCCATCCC
CAGTTCTTAGCCAGTTGAGCCAGCGACAACAGCACCAGAGCCAGGCAGTCACTGTTCTCTCT
CCTGGTTTTGGAGTCCTTTCCCTTCCCAGGCCAAACTTCGAGAATCAACACCTGGAGACAGTCC
CTCCACTGTGAACAAGCTTTTGCAGCTTCCCAGCACGACCATTGAAAATATCTCTGTGTCTG
TCCACCAGCCACAGCCCAAACACATCAAAC TTGCTAAGCGGCGGATACCCCCAGCTTCTAAG
ATCCCAGCTTCTGCAGTGGAATGCCTGGTTCAGCAGATGTCACAGGATTAAATGTGCAGTT
TGGGGCTCTGGAATTTGGGTCAGAACCTTCTCTCTCTGAATTTGGATCAGCTCCAAGCAGTG
AAAATAGTAATCAGATTCCCATCAGCTTGTATTGCAAGTCTTTAAGTGAGCCTTTGAATACA
TCTTTATCAATGACCAGTGCAGTACAGAACTCCACATATACAAC TCCGTCATTACCTCCTG
CAGTCTGACAAGCTCATCACTGAATTCTGCTAGTCCAGTAGCAATGTCTTCCTCTTATGACC
AGAGTTCTGTGCATAACAGGATCCCATACCAAAGCCCTGTGAGTTCATCAGAGTCAGCTCCA
GGAACCATCATGAATGGACATGGTGGTGGTCGAAGTCAGCAGACACTAGACAGTAAGTATAG
CAGCAAGCTACTCTTGTCATGGCTGGTGCCAACCAAACAGAGGAAGAGGATAGCTCACGTGA
TGTGGAAAACACCAGTTGGTCAATGGCTCATTTCGT**TAAA**AAGCAGCCCTTTTGCTTTTTTGT
TTTTGGACCAGGTGTTGGCTGTGGTGTATTAGAAATGTCTTAACCACAGCAAGAAGGAGGT
GGTGGTCTCATATTCTTCTGCCCTAATCAGACTGCACCACAAGTGCAGCATAACAGTATGCAT
TTTAAAGATGCTTGGGCCAGGCGGGGTGGCTGATGCCCATAATCCCAGTGCTTTGGGGGGCC
AAGGCAGGCAGATTGCCCAAGCTCAGGAGTTTGAGACCACCCTGGGCAACATGGTGAAACTC
TGTCTCTACTAAAATACGAAAAACTAGCCGGGTGTGGTGGCGGCGCGTGCCTGTAATCCCAG
CTACTTGGGAGGCTGAGGCACAAGAATCGCTTGAGCCAGCTTGGGCTACAAAGTGAGACTCC
GTCTGAAAAGA

FIGURE 44

MCFKALGRNSVLLRICSFIPLLKSSVLGSGFGELAPPKMANITSSQILDQLKAPSLGQFTTT
PSTQQNSTSHPTTTTTSWDLKPPTSQSSVLSHLDFKSQPEPSPVLSQLSQRQQHQSQAVTVPP
PGLESFPSQAKLRESTPGDSPSTVNKLLQLPSTTIENISVSVHQPPKHIKLAKRRIPPASK
IPASAVEMPGSADVTGLNVQFGALEFGSEPSLSEFGSAPSSSENSNQIPISLYSKSLSEPLNT
SLSMTSAVQNSTYTTTSVITSCSLTSSSLNSASPVAMSSSYDQSSVHNRIPIYQSPVSSSESAP
GTIMNGHGGGRSQQTLDISKYSSKLLLSWLVP TKQRKRIAHVMWKT PVGQWLIR

Signal peptide:

amino acids 1-24

GCCGAGTGGGACAAAGCCTGGGGCTGGGCGGGGGGCC**ATG**GCGCTGCCATCCCGAATCCTGCT
TTGGAAACTTGTGCTTCTGCAGAGCTCTGCTGTTCTCCTGCACTCAGCGGTGGAGGAGACGG
ACGCGGGGCTGTACACCTGCAACCTGCACCATCACTACTGCCACCTCTACGAGAGCCTGGCC
GTCCGCCTGGAGGTCAACGACGGCCCCCGGCCACCCCCGCCTACTGGGACGGCGAGAAGGA
GGTGCTGGCGGTGGCGCGCGGCGCACCCGCGCTTCTGACCTGCGTGAACCGCGGGCACGTGT
GGACCGACCGGCACGTGGAGGAGGCTCAACAGGTGGTGCCTGGGACCGGCAGCCGCCCGGG
GTCCCGCACGACCGCGCGGACCGCCTGCTGGACCTCTACGCGTCGGGCGAGCGCCGCGCCTA
CGGGCCCCCTTTTTCTGCGCGACCGCGTGGCTGTGGGCGCGGATGCCTTTGAGCGCGGTGACT
TCTCACTGCGTATCGAGCCGCTGGAGGTCGCCGACGAGGGCACCTACTCCTGCCACCTGCAC
CACCATTACTGTGGCCTGCACGAACGCCGCGTCTTCCACCTGACGGTCGCCGAACCCACGC
GGAGCCGCCCCCCCCGGGGCTCTCCGGGCAACGGCTCCAGCCACAGCGGCGCCCCAGGCCCAG
ACCCACACTGGCGCGCGGCCACAACGTCATCAATGTCATCGTCCCCGAGAGCCGAGCCAC
TTCTTCCAGCAGCTGGGCTACGTGCTGGCCACGCTGCTGCTCTTCATCCTGCTACTGGTCAC
TGTCCTCCTGGCCGCCCGCAGGCGCCGCGGAGGCTACGAATACTCGGACCAGAAGTCGGGAA
AGTCAAAGGGGAAGGATGTAACTTGGCGGAGTTCGCTGTGGCTGCAGGGGACCAGATGCTT
TACAGGAGTGAGGACATCCAGCTAGATTACAAAAACAACATCCTGAAGGAGAGGGCGGAGCT
GGCCACAGCCCCCTGCCTGCCAAGTACATCGACCTAGACAAAGGGTTCCGGAAGGAGAACT
GCAAA**TAG**GGAGGCCCTGGGCTCCTGGCTGGGCCAGCAGCTGCACCTCTCCTGTCTGTGCTC
CTCGGGGCATCTCCTGATGCTCCGGGGCTCACCCCCCTTCCAGCGGCTGGTCCCGCTTTCCT
GGAATTTGGCCTGGGCGTATGCAGAGGCCGCTCCACACCCCTCCCCAGGGGCTTGGTGGC
AGCATAGCCCCACCCCTGCGGCCTTTGCTCACGGGTGGCCCTGCCACCCCTGGCACAACC
AAAATCCCCTGATGCCCATCATGCCCTCAGACCCTTCTGGGCTCTGCCCGCTGGGGGCCTG
AAGACATTCTGGAGGACACTCCCATCAGAACCTGGCAGCCCCAAACTGGGGTCAGCCTCA
GGGCAGGAGTCCCACTCCTCCAGGGCTCTGCTCGTCCGGGGCTGGGAGATGTTCTGGAGGA
GGACACTCCCATCAGAACTTGGCAGCCTTGAAGTTGGGGTCAGCCTCGGCAGGAGTCCCACT
CCTCCTGGGGTGTGCTGCCCTGCCACCAAGAGCTCCCCACCTGTACCACCATGTGGGACTCCAG
GCACCATCTGTTCTCCCCAGGGACCTGCTGACTTGAATGCCAGCCCTTGCTCCTCTGTGTTG
CTTTGGGCCACCTGGGGCTGCACCCCTGCCCTTTCTCTGCCCCATCCCTACCCTAGCCTTG
CTCTCAGCCACCTTGATAGTCACTGGGCTCCCTGTGACTTCTGACCCTGACACCCCTCCCTT
GGACTCTGCCTGGGCTGGAGTCTAGGGCTGGGGCTACATTTGGCTTCTGTACTGGCTGAGGA
CAGGGGAGGGAGTGAAGTTGGTTTGGGGTGGCCTGTGTTGCCACTCTCAGCACCCACATTT
GCATCTGCTGGTGGACCTGCCACCATCACAATAAAGTCCCCATCTGATTTTTTAAAAAAAAAA
AAAAA

FIGURE 46

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60618

<subunit 1 of 1, 341 aa, 1 stop

<MW: 38070, pI: 6.88, NX(S/T): 1

MALPSRILLWKLVLQSSAVLLHSAVEETDAGLYTCNLHHHYCHLYESLAVRLEVTDGPPAT
PAYWDGEKEVLAVARGAPALLTCVNRGHVWTDHRHVEEAQQVVHWDRQPPGVPHDRADRLDL
YASGERRAYGPLFLRDRVAVGADAFERGFSLRIEPLVADEGTYSCHLHHHYCGLHERRVF
HLTVAEPHAEPPPRGSPGNGSSHSGAPGPDPTLARGHNVINVIVPESRAHFFQQLGYVLATL
LLFILLLVTVLLAARRRRRGYEYSDQKSGKSKGKDVNLAEFAVAAGDQMLYRSEDIQLDYKN
NILKERAELAHSPLPAKYIDLDKGFRKENCK

Important features:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 237-262

N-glycosylation site.

amino acids 205-208

Cell attachment sequence.

amino acids 151-154

Coproporphyrinogen III oxidase proteins.

amino acids 115-140

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 47

CGCCGGAGGCAGCGGCGGCGTGGCGCAGCGGCGAC**ATG**GCCGTTGTCTCAGAGGACGACTTT
CAGCACAGTTCAAACCTCCACCTACGGAACCACAAGCAGCAGTCTCCGAGCTGACCAGGAGGC
ACTGCTTGAGAAGCTGCTGGACCGCCCGCCCCCTGGCCTGCAGAGGCCCGAGGACCGCTTCT
GTGGCACATACATCATCTTCTTCAGCCTGGGCATTGGCAGTCTACTGCCATGGAACCTTCTTT
ATCACTGCCAAGGAGTACTGGATGTTCAAACCTCCGCAACTCCTCCAGCCCAGCCACCGGGGA
GGACCCTGAGGGCTCAGACATCCTGAACTACTTTGAGAGCTACCTTGCCGTTGCCTCCACCG
TGCCCTCCATGCTGTGCCTGGTGGCCAACTTCTGCTTGTCAACAGGGTTGCAGTCCACATC
CGTGTCTTGGCCTCACTGACGGTCATCCTGGCCATCTTCATGGTGATAACTGCACTGGTGAA
GGTGGACACTTCTCCTGGACCCGTGGTTTTTTTTGCGGTCACCATTTGTCTGCATGGTGATCC
TCAGCGGTGCCTCCACTGTCTTCAGCAGCAGCATCTACGGCATGACCGGCTCCTTTCTCTATG
AGGAACTCCCAAGCACTGATATCAGGAGGAGCCATGGGCGGGACGGTCAGCGCCGTGGCCTC
ATTGGTGGACTTGGCTGCATCCAGTGATGTGAGGAACAGCGCCCTGGCCTTCTTCTGACGG
CCACCATCTTCTCGTGTCTGCATGGGACTCTACCTGCTGCTGTCCAGGCTGGAGTATGCC
AGGTACTACATGAGGCCTGTTCTTGCGGCCCATGTGTTTTCTGGTGAAGAGGAGCTTCCCCA
GGACTCCCTCAGTGCCCTTTCGGTGGCCTCCAGATTGATTCCCACACACCCCTCTCC
GCCCCATCCTGAAGAAGACGGCCAGCCTGGGCTTCTGTGTACCTACGTCTTCTTCATCACC
AGCCTCATCTACCCCGCCGTCTGCACCAACATCGAGTCCCTCAACAAGGGCTCGGGCTCACT
GTGGACCACCAAGTTTTTCATCCCCCTCACTACCTTCTCCTGTACAACCTTTGCTGACCTAT
GTGGCCGGCAGCTCACCGCCTGGATCCAGGTGCCAGGGCCCAACAGCAAGGCGCTCCCAGGG
TTCGTGCTCCTCCGGACCTGCCTCATCCCCCTCTTCGTGCTCTGTAACCTACCAGCCCCGCGT
CCACCTGAAGACTGTGGTCTTCCAGTCCGATGTGTACCCCGCACTCCTCAGCTCCCTGCTGG
GGCTCAGCAACGGCTACCTCAGCACCCCTGGCCCTCCTCTACGGGCCTAAGATTGTGCCCAGG
GAGCTGGCTGAGGCCACGGGAGTGGTGATGTCCTTTTATGTGTGCTTGGGCTTAACACTGGG
CTCAGCCTGCTCTACCTCCTGGTGCACCTCATC**TAGA**AAGGGAGGACACAAGGACATTGGTG
CTTCAGAGCCTTTGAAGATGAGAAGAGAGTGCAGGAGGGCTGGGGGCCATGGAGGAAAGGCC
TAAAGTTTCACTTGGGGACAGAGAGCAGAGCACACTCGGGCCTCATCCCTCCCAAGATGCCA
GTGAGCCACGTCCATGCCCATTCCGTGCAAGGCAGATATTCCAGTCATATTAACAGAACACT
CCTGAGACAGTTGAAGAAGAAATAGCACAAATCAGGGGTACTCCCTTCACAGCTGATGGTTA
ACATTCCACCTTCTTTCTAGCCCTTCAAAGATGCTGCCAGTGTTTCGCCCTAGAGTTATTACA
AAGCCAGTGCCAAAACCCAGCCATGGGCTCTTTGCAACCTCCCAGCTGCGCTCATTCCAGCT
GACAGCGAGATGCAAGCAAATGCTCAGCTCTCCTTACCCTGAAGGGGTCTCCCTGGAATGGA
AGTCCCCTGGCATGGTCAGTCCTCAGGCCCAAGACTCAAGTGTGCACAGACCCCTGTGTTCT
GCGGGTGAACAACTGCCCCACTAACCCAGACTGGAAAACCCAGAAAGATGGGCCTTCCATGAAT
GCTTCATTCCAGAGGGACAGAGGGCCTCCCTGTGCAAGGGATCAAGCATGTCTGGCCTGGG
TTTTCAAAAAAAGAGGGATCCTCATGACCTGGTGGTCTATGGCCTGGGTCAAGATGAGGGTC
TTTCAGTGTTCTGTTTACAACATGTCAAAGCCATTGGTTCAAGGGCGTAATAAATACTTGC
GTATTCAAAAA

FIGURE 48

MAVVSEDDFQHSSNSTYGTSSSLRADQEALLEKLLDRPPPGQLQRPEDRFCGTYYIFFSLGI
GSLLPWNFFITAKEYWMFKLRNSSSPATGEDPEGSDILNYFESYLAVASTVPSMLCLVANFL
LVNRVAVHIRVLASLTVILAI FMVITALVKVDTSSWTRGFFAVTIVCMVILSGASTVFSSSI
YGMTGSFPMRNSQALISGGAMGGTVSAVASLVDLAASSDVRNSALAFFLTATIFLVLCMGLY
LLLSRLEYARYYMRPVLA AHVFSGEEELPQDSLSAPSVASRFIDSHTPPLRPILKKTASLGF
CVTYVFFITS LIYPAVCTNIESLNKSGSLWTTKFFIPLTTFLLYNFADLCGRQLTAWIQVP
GPNSKALPGFVLLRTCLIP LFVLCNYQPRVHLKTVVVFQSDVYPALLSSLLGLSNGYLSTLAL
LYGPKIVPRELAEATGVVMSFYVCLGLTLGSACSTLLVHLI

Transmembrane domain:

amino acids 50-74 (type II), 105-127, 135-153, 163-183, 228-252,
305-330, 448-472

MAPLALHLLVLPILLSLVASQDWAERSQDPFEKCMQDPDYEQLLKVVWTWGLNRTLKPQRV
IVVGAGVAGLVAAKVLSDAGHKVTTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRI
HKLCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGALRPQEKGHSPEDIYQMA
LNQALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSSEGGFFYLSFAEALR
AHSCLSDRLQYSRIVGGWDLPLRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNL
KVLKADVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRPPFWREEHIEGG
HSNTDRPSRMIFYPPPREGALLLASYTWSDAAAAFAGLSREEALRLALDDVAALHGPVVRQL
WDGTGVVKRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKS
ALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQ
NTTHTRTSH

amino acids 1-21

FIGURE 51

CTGACATGGCCTGACTCGGGACAGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCCGGCCTTCTGCCTGCAT
 GGACGCTCTGAAGCCACCCTGTCTCTGGAGGAACCACGAGCGAGGGAAGAAGGACAGGGACTCGTGTGGCAGGAA
 GAACTCAGAGCCGGGAAGCCCCCATTCACTAGAAGCACTGAGAGATGCGGCCCCCTCGCAGGGTCTGAATTTCCCT
 GCTGCTGTTACAAAGATGCTTTTTATCTTTAACTTTTTGTTTTCCCACTTCCGACCCCGCGTTGATCTGCAT
 CCTGACATTTGGAGCTGCCATCTTCTTGTGGCTGATCACCAGACCTCAACCCGTCTTACCTCTTCTTGACCTGAA
 CAATCAGTCTGTGGGAATTGAGGGAGGAGCAGGAAGGGGGTTTCCCAGAAGAACAATGACCTAACAAAGTTGCTG
 CTTCTCAGATGCCAAGACTATGTATGAGGTTTTCCAAAGAGGACTCGCTGTGTCTGACAATGGGCCCTGCTTGGG
 ATATAGAAAACCAACCAGCCCTACAGATGGCTATCTTACAAACAGGTGTCTGATAGAGCAGAGTACCTGGGTTT
 CTGTCTCTTGCATAAAGGTTATAAATCATCACCAGACCAGTTTGTGCGCATCTTTGCTCAGAATAGGCCAGAGTG
 GATCATCTCCGAATTGGCTTGTTACACGTAATCTATGGTAGCTGTACCTCTGTATGACACCTTGGGACCAGAAGC
 CATCGTACATATTGTCAACAAGGCTGATATCGCCATGGTGATCTGTGACACACCCCAAAAGGCATTGGTGCTGAT
 AGGGAATGTAGAGAAAGGCTTCAACCCGAGCCTGAAGGTGATCATCCTTATGGACCCCTTGTATGATGACCTGAA
 GCAAAGAGGGGAGAAGAGTGGAATTGAGATCTTATCCCTATATGATGCTGAGAACCTAGGCAAAGAGCACTTCAG
 AAAACCTGTGCCTCTAGCCCAGAAGACCTGAGCGTCATCTGCTTACCAGTGGGACCACAGGTGACCCCAAGG
 AGCCATGATAACCCATCAAAATATTGTTTCAAATGCTGCTGCCCTTCTCAAATGTGTGGAGCATGCTTATGAGCC
 CACTCCTGATGATGTGGCCATATCCTACCTCCCTCTGGCTCATATGTTTGAGAGGATTGTACAGGCTGTTGTGTA
 CAGCTGTGGAGCCAGAGTTGGATTCTTCCAAGGGGATATTCCGTTGCTGGCTGACGACATGAAGACTTTGAAGCC
 CACATTGTTTTCCCGCGGTGCCTCGACTCCTTAACAGGATCTACGATAAGGTACAAAATGAGGCCAAGACACCCCTT
 GAAGAAGTTCTTGTTGAAGCTGGCTGTTTCCAGTAAATTCAAAGAGCTTCAAAGGGTATCATCAGGCATGATAG
 TTTCTGGGACAAGCTCATCTTTGCAAAGATCCAGGACAGCCTGGGCGGAAGGGTTCGTGTAATTGTCACCTGGAGC
 TGCCCCCATGTCCACTTCAGTCATGACATTCTTCCGGGCAGCAATGGGATGTCAAGTGTATGAAGCTTATGGTCA
 AACAGAATGCACAGGTGGCTGTACATTTACATTACCTGGGGACTGGACATCAGGTACGTTGGGGTCCCTTGGC
 TTGCAATTACGTGAAGCTGGAAGATGTGGCTGACATGAACACTTTTACAGTGAATAATGAAGGAGAGGTCTGCAT
 CAAGGGTACAAACGTGTTCAAAGGATACCTGAAGGACCCTGAGAAGACACAGGAAGCCCTGGACAGTGATGGCTG
 GCTTCACACAGGAGACATTGGTCGCTGGCTCCCGAATGGAACCTCTGAAGATCATCGACCGTAAAAAGAACATTTT
 CAAGCTGGCCCCAAGGAGAATACATTGCACCAGAGAAGATAGAAAATATCTACAACAGGAGTCAACCAGTGTTACA
 AATTTTTGTACACGGGGAGAGCTTACGGTCATCCTTAGTAGGAGTGGTGGTTCCTGACACAGATGTACTTCCCTC
 ATTTGCAGCCAAGCTTGGGGTGAAGGGCTCCTTTGAGGAACTGTGCCAAAACCAAGTTGTAAGGGAAGCCATTTT
 AGAAGACTTGCAGAAAATTGGGAAAGAAAGTGGCCTTAAACCTTTTGAACAGGTCAAAGCCATTTTTCTTCATCC
 AGAGCCATTTTTCCATTGAAAATGGGCTCTTGACACCAACATTGAAAGCAAAGCGAGGAGAGCTTTCCAAATACTT
 TCGGACCCAAATTGACAGCCTGTATGAGCACATCCAGGATTAGGATAAGGTACTTAAGTACCTGCCGGCCCACTG
 TGCATGCTTGTGAGAAAATGGATTAAAACTATTCTTACATTTGTTTTGCCTTTCCTCCTATTTTTTTTTTAACC
 TGTTAAACTCTAAAGCCATAGCTTTTTGTTTTATATTGAGACATATAATGTGTAAACTTAGTTCCCAAATAAATCA
 ATCCTGTCTTTCCCATCTTCGATGTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCCTGTCTTCAA
 GATCCCAGTTTATGTTCTGTGTCCTTCTCATGATTTCCAACCTTAATACTATTAGTAACCACAAGTTCAAGGGT
 CAAAGGGACCCTCTGTGCCTTCTTCTTTGTTTTGTGATAAACATAACTTGCCAACAGTCTCTATGCTTATTTACA
 TCTTCTACTGTTCAAACCTAAGAGATTTTTAAATTTCTGAAAACTGCTTACAATTCATGTTTTCTAGCCACTCCAC
 AAACCACTAAAATTTTAGTTTTAGCCTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATGCTCTT
 CTGCGTAAATTAATTTGTGTACTGAAGGGAAAAGTTTGATCATACCAACATTTCTTAACTCTCTAGTTAGATA
 TCTGACTTGGGAGTATTAATAAATTTGGGTCTATGACATACTGTCCAAAAGGAATGCTGTTCTTAAAGCATTATTTA
 CAGTAGGAACCTGGGGAGTAAATCTGTTCCCTACAGTTTGTGCTGAGCTGGAAGCTGTGGGGGAAGGAGTTGACA
 GGTGGGCCCAGTGAACCTTTCCAGTAAATGAAGCAAGCACTGAATAAAAACCTCCTGAACTGGGAACAAAGATCT
 ACAGGCAAGCAAGATGCCACACAACAGGCTTATTTCTGTGAAGGAACCAACTGATCTCCCCACCCCTTGGATT
 AGAGTTCTGCTCTACCTTACCCACAGATAACACATGTTGTTTCTACTTGTAATGTAAAGTCTTTAAATAAAC
 TATTACAGATAAAAAA

FIGURE 52

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60775

<subunit 1 of 1, 739 aa, 1 stop

<MW: 82263, pI: 7.55, NX(S/T): 3

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFLLLF TKMLFIEN
 FLFSPLPTPALICILTFGAAIFLWLITRPQPVLP LLDLNNQSVGIEGGARKGVSQKNNDLTS
 CCFSDAKTMYEVFQ RGLAVSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYKSS
 PDQFVGIFAQNRPEWIISELACYTYSMAVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKAL
 VLIGNVEKGFTPSLKVII LMDPFDDDLKQRGEKSGIEILSLYDAENLGKEHFRKPVPPSPED
 LSVICFTSGTTGDPKGAMITHQNIVSNA AAF LKCV EHA YEPTPDDVAISYLPLAHMFERIVQ
 AVVYSCGARVGFFQGD IRL LADD M KTLKPTLFP AVPRLLNRIYDKVQNEAKT PLKKFLLKLA
 VSSKFKE LQKGIIRHDSFWDKLI FAKIQDSLGRVRVIVTGAAPMSTSVMTFFRAAMGCQVY
 EAYGQTECTGGCTFTLPGDWTSGHVG VPLACNYVKLEDVADMNYFTVNNEGEVCIKGTNVFK
 GYLKDPEKTQEALDS DGWLHTGDIGRWLPNGTLKIIDRKKNI FKL AQGEYIAPEKIENIYNR
 SQPVLQIFVHGESLRSSLVGVVVPD TDVLP SF AAKLG VKGSFEELCQNQVVREAILEDLQKI
 GKESGLKTFEQVKAIFLHPEPFSIENGLLTPTL KAKRGELSKYFRTQIDSLYEHIQD

Important features:

Type II transmembrane domain:

amino acids 61-80

Putative AMP-binding domain signature.

amino acids 314-325

N-glycosylation site.

amino acids 102-105, 588-591 and 619-622

GGAGGCGGAGGAGCCGCGGCGGAGCCGGGCGGAGCAGTGAAGGGCCCTAGCGGGGGCCGAGAGCGGG
CCCGGGGCCCCCTAAGCCATTCTCTGAAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAAT
CCGGT**ATG**GACGACTGGAAGCCCAGCCCCCTCATCAAGCCCTTTGGGGCTCGGAAGAAGCGG
AGCTGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGCCCTGCGGAGATTCTGTCA
GACAGGGGGCCGTGCTTTTCTCTGCTGGTGACTGTCAATTGTCAATATCAAGTTGATCCTGGACA
CTCGGCGAGCCATCAGTGAAGCCAATGAAGACCCAGAGCCAGAGCAAGACTATGATGAGGCC
CTAGGCCGCGCTGGAGCCCCCACGGCGCAGAGGCAGTGGTCCCCGGCGGGTCTCTGGACGTAGA
GGTGTATTCAAGTCGCAGCAAAGTATATGTGGCAGTGGATGGCACCACGGTGTCTGGAGGATC
AGGCCCCGGGAGCAGGGCCGGGGCATCCATGTCAATTGTCTCAACCAGGCCACGGGCCACGTG
ATGGCAAAACGTGTGTTTGACACGTACTCACCTCATGAGGATGAGGCCATGGTGTCTATTCTCT
CAACATGGTAGCGCCCCGGCCGAGTGCTCATCTGCACTGTCAAGGATGAGGGCTCCTTCCACC
TCAAGGACACAGCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGC
TGGAGGGACACATGGGCCTTCGTGGGACGAAAAGGAGGTCTCTGTCTTCGGGGAGAAACATTCT
TAAGTCACCTGCCCTCTCTTCTGGGGGGACCCAGTCTCTGTGAAGACAGATGTGCCATTGA
GCTCAGCAGAAGAGGCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCGCCGGCGC
TTCTGCAGCAAAGTTGAGGGCTATGGAAGTGTATGCAGCTGCAAGGACCCACACCCATCGA
GTTTCAGCCCTGACCCACTCCCAGACAACAAGGTCTCAATGTGCCTGTGGCTGTCAATTGCAG
GGAACCGACCCAATTACCTGTACAGGATGCTGCGCTCTCTGCTTTCAGCCCAGGGGGTGTCT
CCTCAGATGATAACAGTTTTTCATTGACGGCTACTATGAGGAACCCATGGATGTGGTGGCACT
GTTTGGTCTGAGGGGCATCCAGCATACTCCCATCAGCATCAAGAATGCCCGCGTGTCTCAGC
ACTACAAGGCCAGCCTCACTGCCACTTTCAACCTGTTTCCGGAGGGCCAAGTTTGCTGTGGTT
CTGGAAGAGGACCTGGACATTGCTGTGGATTTTTTCAGTTTCTGAGCCAATCCATCCACCT
ACTGGAGGAGGATGACAGCCTGTACTGCATCTCTGCCTGGAATGACCAAGGGGTATGAACACA
CGGCTGAGGACCCAGCACTACTGTACCGTGTGGAGACCATGCCTGGCTGGGTGGGTGCTC
AGGAGGTCTTGTACAAGGAGGAGCTTGAGCCCAAGTGGCCTACACCGGAAAAGCTCTGGGA
TTGGGACATGTGGATGCGGATGCCTGAACAACGCCGGGGCCGAGAGTGCATCATCCCTGACG
TTTCCCGATCCTACCACTTTGGCATCGTCGCGCTCAACATGAATGGCTACTTTACGAGGCC
TACTTCAAGAAGCACAAGTTCAACACGGTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCT
GAAGAAAGAAGCTTATGAAGTGGAAGTTCACAGGCTGCTCAGTGAGGCTGAGGTTCTGGACC
ACAGCAAGAACCCTTGTGAAGACTCTTTCTGCCAGACACAGAGGGCCACACCTACGTGGCC
TTTATTTCGAATGGAGAAAGATGATGACTTCACCACCTGGACCCAGCTTGCCAAGTGCCCTCA
TATCTGGGACCTGGATGTGCGTGGCAACCATCGGGGCCCTGTGGAGATTGTTTCGGAAGAAGA
ACCACTTCTCTGGTGGTGGGGGTCCCGGCTTCCCCCTACTCAGTGAAGAAGCCACCCTCAGTC
ACCCCAATTTTCTGGAGCCACCCCCAAAGGAGGAGGGAGCCCCAGGAGCCCCAGAACAGAC
ATGAGACCTCCTCCAGGACCCTGCGGGGCTGGGTACTGTGTACCCCCAGGCTGGCTAGCCCT
TCCCTCCATCCTGTAGGATTTTTGTAGATGCTGGTAGGGGCTGGGGCTACCTTGTTTTTTAACA
TGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCCTGCTCCAACACCCCGTTCTGAGTT
AAAAGTCTATTTATTTACTTCCTTGTGGAGAAGGGCAGGAGAGTACCTGGGAATCATTACG
ATCCCTAGCAGCTCATCCTGCCCTTTGAATACCCTCACTTTCCAGGCCTGGCTCAGAATCTA
ACCTATTTATTGACTGTCCTGAGGGCCTTGAAAAACAGGCCGAACCTGGAGGGCCTGGATTTC
TTTTTGGGCTGGAATGCTGCCCTGAGGGTGGGGCTGGCTCTTACTCAGGAAACTGCTGTGCC
CAACCCATGGACAGGCCCAGCTGGGGCCCATGCTGACACAGACTCACTCAGAGACCCTTA
GACACTGGACCAGGCCTCCTCTCAGCCTTCTCTTTGTCCAGATTTCCAAAGCTGGATAAGTT
GGTCATTGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 54

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA61185

><subunit 1 of 1, 660 aa, 1 stop

><MW: 75220, pI: 6.76, NX(S/T): 0

MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQALRRFCQTGAVLFLLVTVIVNIKLILDTR
 RAISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVLDVEVYSSRSKVYVAVDGTTVLEDEA
 REQGRGIHVIVLNQATGHVMAKRVEDTYSPHEDEAMVLFNLMVAPGRVLICTVKDEGSFHLK
 DTAKALLRSLGSQAGPALGWRDTWAFVGRKGGPVFGEKHSKSPALSSWGDPVLLKTDVPLSS
 AEEAECHWADTELNRNRRRRFCCKVEGYGSVCCKDPTPIEFSPDPLPDNKVLNVPVAVIAGN
 RPNYLYRMLRSLLSAQGVSPQMITVFIDGYEPMDDVVALFGLRGIQHTPISIKNARVSQHY
 KASLTATFNLFPFAKFAVVLEEDLDIAVDFFSFLSQSIHLLEEDDSLYCISAWNDAQYEHTA
 EDPALLYRVETMPGLGWVLRRLSYKEELEPKWPTPEKLWDWDMWMMRMPEQRRGRECIIPDVS
 RSYHFGIVGLNMNGYFHEAYFKKHKFNTVPGVQLRNVDLSLKKEAYEVEVHRLLEAEVLDHS
 KNPCEDSFLPDTEGHTYVAFIRMEKDDDDFTTWTQLAKCLHIWDLDVIRGNHRGLWRLFRKKNH
 FLVVGVPASPYSVKKPPSVTPIFLEPPPKEEGAPGAPEQT

Important features of the protein:

Transmembrane domain:

amino acids 38-55

Homologous region to Mouse GNT1

amino acids 229-660

FIGURE 55

CGGACGCGTGGGCTGCTGGTGGGAAGGCCTAAAGAACTGGAAAGCCCACTCTCTTGGAAACCACCACAC
 CTGTTTAAAGAACCTAAGCACCATTAAAGCCACTGGAAATTTGTTGTCTAGTGGTGTGGGTGAATA
 AAGGAGGGCAGAAATGATGATTTTCATCTCCATTAGCCTGCTGTCTCTGGCTATGTTGGTGGGATGTTA
 CGTGGCCGGAATCATTCCCTTGGCTGTTAATTTCTCAGAGGAACGACTGAAGCTGGTGACTGTTTTGG
 GTGCTGGCCTTCTCTGTGGAAGTGTCTGGCAGTCATCGTGCCTGAAGGAGTACATGCCCTTTATGAA
 GATATTTCTTGAGGGAAAACACCACCAAGCAAGTGAACACATAATGTGATTGCATCAGACAAAGCAGC
 AGAAAAATCAGTTGTCCATGAACATGAGCAGCCACGACACACACAGCTGCATGCCATATATTGGTG
 TTTCCCTCGTTCTGGGCTTCGTTTTTCATGTTGCTGGTGGACCAGATTGGTAACTCCCATGTGCATTCT
 ACTGACGATCCAGAAGCAGCAAGGTCTAGCAATTCCAAAATCACCACCACGCTGGGTCTGGTTGTCCA
 TGCTGCAGCTGATGGTGTGCTTTGGGAGCAGCAGCATCTACTTCACAGACCAGTGTCCAGTTAATTG
 TGTTTGTGGCAATCATGCTACATAAGGCACCAGCTGCTTTTGGACTGGTTTCCTTCTTGATGCATGCT
 GGCTTAGAGCGGAATCGAATCAGAAAGCACTTGTGGTCTTTGCATTGGCAGCACCAGTTATGTCCAT
 GGTGACATACTTAGGACTGAGTAAGAGCAGTAAAGAAGCCCTTTCAGAGGTGAACGCCACGGGAGTGG
 CCATGCTTTTTCTCTGCCGGGACATTTCTTTATGTTGCCACAGTACATGTCCTCCCTGAGGTGGGCGGA
 ATAGGGCAGGCCACAAGCCCCGATGCCACGGGAGGGAGAGGCCTCAGCCGCCTGGAAGTGGCAGCCCT
 GGTCTGGGTTGCCTCATCCCTCTCATCCTGTCTAGTAGGACACCAGCATTAATGTTCAAGGTCCAGC
 CTTGGTCCAGGGCCGTTTGCCATCCAGTGAGAACAGCCGGCACGTGACAGCTACTCACTTCCTCAGTC
 TCTTGTCTCACCTTGCGCATCTCTACATGTATTCTAGAGTCCAGAGGGGAGGTGAGGTTAAAACCTG
 AGTAATGAAAAGCTTTTAGAGTAGAAACACATTTACGTTGCAGTTAGCTATAGACATCCCATTTGTGT
 TATCTTTTAAAGGCCCTTGACATTTTGCGTTTTAATATTTCTCTTAACCTATTCTCAGGGAAGATG
 GAATTTAGTTTTAAGGAAAAGAGGAGAATTCATACTCACAATGAAATAGTGAATGAAAATACAGT
 GTTCTGTAATTAAGCTATGTCTCTTTCTTTAGTTAGAGGCTCTGCTACTTTATCCATTGATTTTT
 AACATGGTTCCCACTATGTAAGAGCTGGTGTCTTTAGCCTCTATGCCACATGCGTTGATGGAAGGTCATA
 GCACCCACTCACTTAGATGCTAAAGGTGATTCTAGTTAATCTGGGATTAGGGTCAGGAAAATGATAGC
 AAGACACATTGAAAGCTCTCTTTATACTCAAAGAGATATCCATTGAAAAGGGATGTCTAGAGGGATT
 TAAACAGCTCCTTTGGCACGTGCCTCTCTGAATCCAGCCTGCCATTCCATCAAATGGAGCAGGAGAGG
 TGGGAGGAGCTTCTAAAGAGGTGACTGGTATTTTGTAGCATTCTTGTCAAGTTCTCCTTTGCAGAAT
 ACCTGTCTCCACATTCTTAGAGAGGAGCCAAGTTCTAGTAGTTTCAGTTCTAGGCTTTCTTCAAGAA
 CAGTCAGATCACAAAGTGTCTTTGGAAATTAAGGGATATTAATTTTAAAGTGATTTTTTGGATGGTTAT
 TGATATCTTTGTAGTAGCTTTTTTTTAAAGACTACCAAATGTATGGTTGTCTTTTTTTTTTGT
 TTTTTTTTTAATTATTTCTCTTAGCAGATCAGCAATCCCTCTAGGGACCTAAATACTAGGTCAGCTTT
 GGCGACACTGTGTCTCTCACATAACCACCTGTAGCAAGATGGATCATAAATGAGAAGTGTTTGCCTA
 TTGATTTAAAGCTTATTGGAATCATGTCTCTTGTCTCTTCGTCTTTTCTTTGCTTTTCTTCTAACTTT
 TCCCTCTAGCCTCTCCTCGCCACAATTTGCTGCTTACTGCTGGTGTTAATATTTGTGTGGGATGAATT
 CTTATCAGGACAACCACTTCTCGAACTGTAATAATGAAGATAATAATATCTTTATTCTTTATCCCCTT
 CAAAGAAATTACCTTTGTGTCAAATGCCGCTTTGTTGAGCCCTTAAAATACCACCTCCTCATGTGTAA
 ATTGACACAATCACTAATCTGGTAATTTAAACAATTGAGATAGCAAAAGTGTTTAAACAGACTAGGATA
 ATTTTTTTTTTCATATTTGCCAAAATTTTGTAAACCCTGTCTTGTCAAATAAGTGTATAATATTGTAT
 TATTAATTTATTTTACTTTCTATACCATTTCAAACACATTACACTAAGGGGGAACCAAGACTAGTT
 TCTTCAGGGCAGTGGACGTAGTAGTTTGTAAAACGTTTCTATGACGCATAAGCTAGCATGCCTATG
 ATTTATTTCTTTCATGAATTTGTCACTGGATCAGCAGCTGTGGAAATAAAGCTTGTGAGCCCTCTGCT
 GGCCACAGTGAGGAAAGTAGCACAAATAGGATACAGTTGTATGTAGTCATTGGCAACAATTGCATACA
 ATTTTACTACCAAGAGAAGGTATAGTATGGAAAGTCCAAATGACTTCCTTGATTGGATGTTAACAGCT
 GACTGGTGTGAGACTTGAGGTTTCATCTAGTCCTTCAAACCTATATGGTTGCCTAGATTCTCTCTGGA
 AACTGACTTTGTCAAATAAATAGCAGATTGTAGTGTCAAAAAA

FIGURE 56

MDDFISISLLSLAMLVGCYVAGIIP LAVNFSEERLKLVTVLGAGLLCGTALAVIVPEGVHAL
YEDILEGKHHQASETHNVIASDKAAEKSVVHEHEHSHDHTQLHAYIGVSLVLGFVFMLLVDQ
IGNSHVHSTDDPEAARSSNSKITTTTLGLVVHAAADGVALGAAASTSQTSVQLIVFVAIMLHK
APAAFGLVSFLMHAGLERNRIRKHL LVFALAAPVMSMVTYLGLSKSSKEALSEVNATGVAML
FSAGTFLYVATVHVLPEVGGIGHSHKPDATGGRGLSRLEVAALVLGCLIP LILSVGHQH

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 37-56, 106-122, 211-230, 240-260, 288-304

GCTCGAGGCGCGGCGGCGGAGAGCGACCCGGCGGCCTCGTAGCGGGGCCGGATCCC
CGAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAGATGATGGGCTTGG
GAAACGGGCGTTCGCAGCATGAAGTCGCCGCCCTCGTGCTGGCCGCCCTGGTGGCCTGCATC
ATCGTCTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGGACCTCCAGACACGGAT
CATGGAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGGCGCCGTGGAGCTGAAGA
AGAACGAGTTCCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGC
CACAACCTTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGGCGGTTTTTGGTGAA
TAACATCACACAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGA
GGAATTACGGCAGGCTGCAGCAGGATGTCCTCCAGTTTTCAGAAGAACCAGACCAACCTGGAG
AGGAAGTTCTCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTG
TGAGGAGCGAATAGAAGAGGTCACCAAAAAGGGGAATGAAGCTGTAGCTTCCAGAGACCTGA
GTGAAAACAACGACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAG
GCAGCAGGCCTGCCACACACAGAGGTGCCACAAGGGAAGGGAACGTGCTTGGTAACAGCAA
GTCCCAGACACCAGCCCCCAGTTCCGAAGTGGTTTTTGGATTCAAAGAGACAAGTTGAGAAAG
AGGAAACCAATGAGATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAG
CCAGGCCGGGAGCAGGTGGTGGAAAGACAGACCTGTAGGTGGAAGAGGCTTCGGGGGAGCCGG
AGAACTGGGCCAGACCCACAGGTGCAGGCTGCCCTGTCAGTGAGCCAGGAAAATCCAGAGA
TGGAGGGCCCTGAGCGAGACCAGCTTGTCATCCCCGACGGACAGGAGGAGGAGCAGGAAGCT
GCCGGGGAAGGGAGAAACCAGCAGAACTGAGAGGAGAAGATGACTACAACATGGATGAAAA
TGAAGCAGAATCTGAGACAGACAAGCAAGCAGCCCTGGCAGGGAATGACAGAAAACATAGATG
TTTTTAATGTTGAAGATCAGAAAAGAGACACCATAAATTTACTTGATCAGCGTGAAAAGCGG
AATCATACACTCTTGAATTGAACTGGAATCACATATTTCAACAGGGCCGAAGAGATGACTA
TAAATGTTTCATGAGGGACTGAATACTGAAAACGTGTAATGTACTAAATAAAATGTACATCTGA

FIGURE 58

MMGLGNRRSMKSPPLVLAALVACIIVLGFNYWIIASSRSVDLQTRIMELEGRVRRAAAERGA
VELKKNEFQGELEKQREQLDKIQSSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQL
KTLQRNYGRLQQDVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVA
SRDLSENNDQRQQLQALSEPQPRQLQAAGLPHTTEVPQGKGNVLGNSKSQTPAPSSEVVLDSCR
QVEKEETNEIQVVNEEPQRDRLPQEPGREQVVEDRPVGGRGFGGAGELGQTPQVQAALSVSQ
ENPEMEGPERDQLVIPDGQEEEQEAAGEGRNQKLRGEDDYNMDENEAESETDKQAALAGND
RNIDVFNVEDQKRDTINLLDQREKRNHTL

Signal peptide:

amino acids 1-29

Downloaded from www.sciencedirect.com

GGATGTCAGAAAGCCTCAGTGTGTGCTCTTCTCTGGCCCTGGGTCTGCTTCTCTTCTACGCTGGCAATTGCCCTCTTCA
CCAGTGGCTTCTCTGCTCACCCTGTTTGGAGCTCACCACCATAGCAGCTGCCAAGAGCCCCCAGGCCCTGGGTCCC
TGCCATGGGGGAGCCAAGGGAAACCTGGGGCCTGCTGGATGGCTTCCCGATTTTCGCGGGTTGTGTTGGTGTCTGA
TAGATGGCTCTGCGATTTGACTTTCGCCACGCCACCCAGCATTCACACGTGCCATAGAGAGCCTCTGTCTCCCTACCC
TCTGTGCGAAACATTAAGCTCCTTGCAGAGGATCCTGGAGATTACAGCCCCACCATGCCCGGCTCTACCGATCTCAGG
TTGACCTCTTACCACCACCATGCAGCGCCTCAAGGCCCTCACCCTGGCTACTGCCTACCTTTATTTGATGCTGTG
GTAGTAACCTTCGCCAGCCACGCCATAGTGAAGACAATCTCATTAAGCAGCTCACCAGTGCAGGAAGGCGTGTAG
TCTTCATGGGAGATGATACCTGGAAGACCTTTTCCCTGGTGTCTTCTCCAAAGCTTTCTTCTTCCCATCCTTCA
ATGTGCAGAGACCTAGACACAGTGGACAATGGCATCCTGGAACACCTCTACCCACCACCATGGACAGTGGTGAATGGG
ACGTGCTGATTTGCTCACTTCTCTGGGTGTGGACCACTGTGGCCACAAGCATGGCCCTCACCACCTGAAATGGCCA
AGAAACTTAGCCAGATGGACCAGGTGATCCAGGACTTGTGGAGCGCTGGAGGAATGACACACTGCTGGTAGTGG
CTGGGGACCATGGGATGACCACAAATGGAGACCATGGAGGGGACAGTGAGCTGGAGGTCTCAGCTGCTCTCTTTTC
TGTATAGCCCCACAGCAGTCTTCCCCAGCACCCCCACCAGAGGAGCCAGAGGTGATTCTCAAGTTAGCCTTGTGC
CCACGCTGGCCCTGCTGCTGGGCCTGCCATCCCATTGGGAATATCGGGGAAGTGATGGCTGAGCTATTCTCAG
GGGGTGAGGACTCCCCAGCCCCACTCCTCTGCTTTTAGCCCAAGCCTCAGCTCTCCATCTCAATGCTCAGCAGGTGT
CCCCGATTTCTTACATCTACTCAGCTGCTACTCAGGACCTTCAAGCTAAGGAGCTTCTCAGCTGCAGAACCTCT
TCTCCAAGGCCCTCTGCTGACTACCAGTGGCTTCTCCAGAGCCCCAAGGGGGCTGAGGCGACACTGCCACTGTGA
TTGCTGAGCTGCAGCAGTTCTTGCGGGGAGCTCGGGCCATGTGCATCGAGTCTTGGGCTCGTTTCTCTCTGGTCC
GCATGGCGGGGGGTACTGCTCTCTTGGCTGCTTCTGCTTTATCTGCCTGCTGGCATCTCAGTGGGCAATATCCC
CAGGCTTTCCATTCTGCCCTCTACTCCTGACACCTGTGGCCTGGGGCCTGGTTGGGGCCATAGCGTATGCTGGAC
TCTTGGGAACATTTGAGCTGAAGCTAGATCTAGTGCTTCTAGGGCTGTGGCTGCAGTGAGCTCATTCTCCCTT
TCTGTGGGAAGCGTGGGCTGGCTGGGGGTCCAAGAGGCCCTGGCAACCCTGTTTCCCATCCCTGGGCGCGTCC
TGTTACTCCTGCTGTTTTCGCTTGGCTGTGTTCTTCTCTGATAGTTTGTGTGTAGCTGAGGCCAGGGCCACCCCT
TCTTTTGGGCTCATTCATCCTGCTCCTGGTTGTCCAGCTTCACTGGGAGGGCCAGCTGCTTCCACCTAAGCTAC
TCACAATGCCCCGCTTGGCACTTCAGCCACAACAACCCCCACGGCACAATGGTGCATATGCCCTGAGGCTTG
GAATTGGGTTGCTTTTATGTACAAGGCTAGCTGGGCTTTTCTACGTTGCCCTGAAGAGACACCTGTTTGGCACT
CCTCTCCCTGGCTGAGTCTCTCGGCATCCATGGTGGTGGTGCAGCCAGAATTTATGGTATGGAGCTTGTGTGG
CGGCGCTGGTGCCCTGTTAGCTGCCGTGCGCTTGTGGCTTCGCGCTATGGTAATCTCAAGAGCCCCGAGCCAC
CCATGCTCTTTGTGCGCTGGGGACTGCCCTAATGGCATTTGGGTACTGCTGCCCTACTGGGCATTGGCGTCGGGG
CAGATGAGGCTCCCCCGCTCTCCGGGTCTGGTCTCTGGGGCATCCATGGTGTGCCTCGGGCTGTAGCAGGGC
TGGCTGCTTTCAGGGCTCGCGCTGCTGCTCTGGAAGCCTGTGACAGTGCTGGTGAAGGCTGGGGCAGGCGCTCCAA
GGACCAGGACTGTCTCTACTCCCTTCTCAGGGCCCCCACTTCTCAAGCTGACTTGGATTATGTGCTCCCTCAA
TCTACCGACACATGCAGGAGGAGTTCGGGGCCCGTTAGAGGAGCAAAATCTCAGGGTCCCCTGACTGTGGCTG
CTTATCAGTTGGGGAGTGTCTACTCAGCTGCTATGGTCACAGCCCTCACCCTGTTGGCCTTCCCACCTTCTGCTGT
TGCATGCGGAGCGCATCAGCCTTGTGTTTCTGCTTCTGTTTCTGCAGAGCTTCTTCTCCTACATCTGCTTGTCTG
CTGGGATACCCGTCACCACCCCTGGTCTTTTACTGTGCCATGGCAGGCAGTCTCGGCTTGGGCCCCATAGGCCA
CAGAGACCTTCTACTCTCACAGGCCACAGCCTGTCTTTCAGGCCATCCATTGGCATGCAGCCTTCTGTGGGATTC
CAGAGGGTCATGGCTCCTGTACTTGGCTGCCGTGCTTGTGCTAGTGGGAGCCAACACCTTGGCTCCCACCTCCTCT
TTGCAGTAGGTTGCCCACTGCTCCTGCTCTGGCCTTTCTGTGTGAGAGTCAAGGGCTGCGGAAGAGACAGCAGC
CCCCAGGGAATGAAGCTGATGCCAGAGTACAGCCCAGGAGGAAGAGGAGCCACTGATGGAGATGCGGCTCCGGG
ATGCGCCTCAGCAGCTTCTATGTCAGCACTGCTGCAGCTGGGCTCAAGTACCTCTTTATCCTTGGTATTCAGATT
TGGCCTGTGCTTGGCAGCCTCCATCCTTCCGAGGCATCTCATGGTCTGGAAGTGTTTGCCCTTAAGTTTCATAT
TTGAGGCTGTGGGCTTCATTGTGAGCAGCGTGGGACTTCTCCTGGGCATAGACTTTGGTGATGAGAGTGGATGGTG
CTGTGAGCTCCTGGTTACGGCAGCTATTTCTGGCCCAGCAGAGGTAGCCTAGTCTGTGATTACTGGCACTTGGCT
ACAGAGAGTGCTGGAGAACAGTGTAGCCTGGCCTGTACAGGTACTGGATGATCTGCAAGACAGGCTCAGCCATAC
TCTTACTATCATGCAGCCAGGGGCGCTGACATCTAGGACTTCATTATTTCTATAAATCAGGACCACAGTGGAGTA
TGATCCCTAACTCCTGATTTGGATGCATCTGAGGGAACAAGGGGGGCGGTCTCCGAAGTGAATAAAATAGGCCGG
CGTGGTGACTTGCACCTATAATCCAGCACTTTGGGAGGCAGAGTGGGAGGATGCTTGGTCCAGGAGTTCA
AGACCAGCCTGTGGAACATAACAAGACCCCGTCTCTACTATTTAAAAAAAAGTGAATAAAATGTAATAAT

FIGURE 60

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62809

<subunit 1 of 1, 1089 aa, 1 stop

<MW: 118699, pI: 8.49, NX(S/T): 2

MQKASVLLFLAWVCFLFYAGIALEFTSGFLLTRLELTNHSSCQEPPGPGSLPWGSQGKPGACW
 MASRFSRVVLVLIDALRFDFAPQHQSHVPREPPVSLPFLGKLSSLQRILEIQPHHARLYRSQ
 VDPPTTTMQRLKALTGSLPTFIDAGSNFASHAIVEDNLIKQLTSAGRRVVFMDGDDTWKDLF
 PGAFSKAFFFPFNFVRDLDTVDNGILEHLYPTMDSGEWDVLIHFLGVDHCGHKHGPHHP
 AKKLSQMDQVIQGLVERLENDTLLVVAGDHGMTTNGDHGGDSELEVSAALFLYSPTAVFPST
 PPEEPEVIPQVSLVPTLALLLGLPIPFNGIENVMAELFSGGEDSQPHSSALAAQASALHLNAQ
 QVSRFLHTYSAATQDLQAKELHQLQNLFSKASADYQWLLQSPKGAEATLPTVIAELQQFLRG
 ARAMCIESWARFSLVRMAGGTALLAASCFCILLASQWAI SPGFPCPLLLTPVAWGLVGAIA
 YAGLLGTIELKLDLVLLGAVAAVSSFLPFLWKAWAGWGSKRPLATLFPPIPGPVLLLLLFRLA
 VFFSDSFVVAEARATPFLGSLFILLVQVQLHWEGQLLPKLLTMPRLGTSATTNPPRHNGAY
 ALRLGIGLLLCTRLAGLFHRCPEETPVCHSSPWLSPLASMVGGRAKNLWYGACVAALVALLA
 AVRLWLRRYGNLKSPEPPMLFVRWGLPLMALGTAAYWALASGADEAPPRLRVLVSGASMLVP
 RAVAGLAASGLALLLWKPVTVLVKAGAGAPRTRTVLTPFSGPPTSQADLDYVVPQIYRHMQE
 EFRGRLERTKSQGPLTVAAYQLGSVYSAAMVTALTLLAFPLLLLHAERISLVFLLLFLQSFL
 LLHLLAAGIPVTTTPGPFTVPWQAVSAWALMATQTFYSTGHQPVFPAIHWHA AFVGFPEGHGS
 CTWLPALLVGANTFASHLLFAVGCPLLLLWPFLCESQGLRKRQQPPGNEADARVRPEEEEEEP
 LMEMRLRDAPQHFYAALLQLGLKYLFI LGIQLACALAASILRRHLMVWKVFAPKFI FEAVG
 FIVSSVGLLLGIALVMRVDGAVSSWFRQLFLAQQR

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domains:

amino acids 317-341, 451-470, 481-500, 510-527, 538-555, 831-850,
 1016-1034, 1052-1070

Leucine zipper pattern.

amino acids 843-864

N-glycosylation sites.

amino acids 37-40, 268-271

FIGURE 61

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGT
GTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA**ATGT**
CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATT
TCATCACAAATTGGCCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGT
CCTGACAGCTCCAGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAA
TATACTCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCC
CAGTGTGTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGT
ACACGTGGAGTCCTTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTG
CCAGGACTTTGAAAGATCAATCATCAGAGTTC AAGGCTAAAATCATCTTCTGGTATGTTTTG
CCCATATCTATTACCGTGTTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCA
CGTTGGCAAAGAGAAACACCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAA
GATTCCTTTGTGCCTGCTGAAAAAATCGTGATTAACCTTTATCACCCCTCAATATCTCGGATGAT
TCTAAAATTTCTCATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAA
TGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAG
GGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCT
CTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAAACAGTCATTGAATATGA
ATATGATGTCAGAACCACTGACATTTGTGCGGGGCTGAAGAGCAGGAGCTCAGTTTGCAGG
AGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCC
CAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCA
CACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCC
AAACTGGCAGGCTGTGTATTCCTTCGCTGTGECAGCTTCGACCAGGATTCAGAGGGCTGCGAG
CCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGC
TCCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGT
TATATGTGCAGATGGAAAAC**TGA**TGCCAACACTTCCTTTTGCCTTTTGTTTCCTGTGCAAC
AAGTGAGTCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGT
TTGTGAGTGTCTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTG
GTTTCATGCATGTAGGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGT
TGTTCTATGCAGAGAAAGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGG
TGGGTGT

FIGURE 62

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62815

<subunit 1 of 1, 442 aa, 1 stop

<MW: 49932, pI: 4.55, NX(S/T): 5

MSYNGHLHQRVFKELKLLTLCSISSQIGPPEVALTTDEKSISSVVLTAPEKWKRNPEDLPVSMQ
QIYSNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQ
CARTLKDQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNED
KRFFVPAEKIVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKH
LGYASHLMEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEDVRTTDICAGPEEQELSL
QEEVSTQGTLLESQAALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWD
PQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEW
GLYVQMEN

Important features:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 63

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCG
TCTGCC**ATG**GGGCTCGGGTTGAGGGGCTGGGGACGTCTCTGCTGACTGTGGCCACCGCCCT
GATGCTGCCCCGTGAAGCCCCCGCAGGCTCCTGGGGGGCCCAGATCATCGGGGGCCACGAGG
TGACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCTTCGGGGGCCAACATCACTGCGGA
GGCTTCCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCACAGAGACCT
CCGCACTGGCCTGGTGGTGTGCTGGGCGCCACGTCTGAGTACTGCGGAGCCCACCCAGCAGG
TGTTTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGAC
ATCTGCCTGCTGCGGCTGAACGGCTCTGCTGTCTGGGCCCTGCAGTGGGGCTGCTGAGGCT
GCCAGGGAGAAGGGCCAGGCCCCCACAGCGGGGACACGGTGCCGGGTGGCTGGCTGGGGCT
TCGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGGCCAAGGTCCGAGTGCTGGAC
CCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCACCCGCAG
TGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGGCCCTGGTGTGCAGGA
ACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCGAC
GTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATCTGGGACGTGGTTCGGCGGAGCAGTCC
CCAGCCCGGCCCCCTGCCTGGGACCACCAGGCCCCCAGGAGAAGCCGCC**TGAG**CCACAACCT
TGCGGCATGCAAATGAGATGGCCGCTCCAGGCCTGGAATGTTCCGTGGCTGGGCCCCACGGG
AAGCCTGATGTTTCAAGGTTGGGGTGGGACGGGCAGCGGTGGGGCACACCCATTCCACATGCA
AAGGGCAGAAGCAAACCCAGTAAAATGTAACTGACAAAAAAAAAAAAAAAAAAAAAGAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 64

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62845

><subunit 1 of 1, 283 aa, 1 stop

><MW: 30350, pI: 9.66, NX(S/T): 2

MGLGLRGWGRPLLTVATALMLPVKPPAGSWGAIIGGHEVTPHSRPYMASVRFGGQHHCGGF
LLRARWVVSAAHCFSHRDLRTGLVVLGAHVLSTAEP TQQVFGIDALTTHPDYHPMTHANDIC
LLRLNGSAVLGPAVGLLRLLPGRRARPPPTAGTRCRVAGWGFVSDFEELPPGLMEAKVRVLDPD
VCNSSWKGHLLTLMCTRS GDSHRRGFCSADSGGPLVCRNRAHGLVSFSGLWCGDPKTPDVY
TQVSAFVAWIWDVVRSSPQPGPLPGTTTRPPGEAA

Signal peptide:

amino acids 1-30

64/249

FIGURE 65

GAGCTACCCAGGCGGCTGGTGTGCAGCAAGCTCCGCGCCGACTCCGGACGCCTGACGCCTGA
 CGCCTGTCCCCGGCCCGGC**ATG**AGCCGCTACCTGCTGCCGCTGTCGGCGCTGGGCACGGTAG
 CAGGCGCCCGCGTGCTGCTCAAGGACTATGTCAACGGTGGGGCTTGCCCCAGCAAGGCCACC
 ATCCCTGGGAAGACGGTCATCGTGACGGGCGCCAACACAGGCATCGGGAAGCAGACCGCCTT
 GGAAGTGGCCAGGAGAGGAGGCAACATCATCCTGGCCTGCCGAGACATGGAGAAGTGTGAGG
 CGGCAGCAAAGGACATCCGCGGGGAGACCCTCAATCACCATGTCAACGCCCGGCACCTGGAC
 TTGGCTTCCCTCAAGTCTATCCGAGAGTTTGCAGCAAAGATCATTGAAGAGGAGGAGCGAGT
 GGACATTCTAATCAACAACGCGGGTGTGATGCGGTGCCCCACTGGACCACCGAGGACGGCT
 TCGAGATGCAGTTTGGCGTTAACCACCTGGGTCACTTTCTCTTGACAACTTGCTGCTGGAC
 AAGCTGAAAGCCTCAGCCCCCTTCGCGGATCATCAACCTCTCGTCCCTGGCCCATGTTGCTGG
 GCACATAGACTTTGACGACTTGAAGTGGCAGACGAGGAAGTATAACACCAAAGCCGCCTACT
 GCCAGAGCAAGCTCGCCATCGTCCTCTTCACCAAGGAGCTGAGCCGGCGGGCTGCAAGGCTCT
 GGTGTGACTGTCAACGCCCTGCACCCCGGCGTGGCCAGGACAGAGCTGGGCAGACACACGGG
 CATCCATGGCTCCACCTTCTCCAGCACCACTCGGGCCCCTCTTCTGGCTGCTGGTCAAGA
 GCCCCGAGCTGGCCGCCAGCCCAGCACATACTGGCCGTGGCGGAGGAACTGGCGGATGTT
 TCCGGAAGTACTTCGATGGACTCAAACAGAAGGCCCGCCCCGAGGCTGAGGATGAGGA
 GGTGGCCCCGAGGCTTTGGGCTGAAAGTGCCCGCCTGGTGGGCTTAGAGGCTCCCTCTGTGA
 GGGAGCAGCCCCCTCCCCAGAT**TAA**CCTCTGGAGCAGATTTGAAAGCCAGGATGGCGCCTCCAG
 ACCGAGGACAGCTGTCCGCCATGCCCGCAGCTTCCTGGCACTACCTGAGCCGGGAGACCCAG
 GACTGGCGGCCGCCATGCCCGCAGTAGGTTCTAGGGGGCGGTGCTGGCCGCAGTGGACTGGC
 CTGCAGGTGAGCACTGCCCCGGGCTCTGGCTGGTTCCGTCTGCTCTGCTGCCAGCAGGGGAG
 AGGGGCCATCTGATGCTTCCCCTGGGAATCTAAACTGGGAATGGCCGAGGAGGAAGGGGCTC
 TGTGCACTTGCAGGCCACGTCAGGAGAGCCAGCGGTGCCTGTGCGGGAGGGTTCCAAGGTGC
 TCCGTGAAGAGCATGGGCAAGTTGTCTGACACTTGGTGGATTCTTGGGTCCCTGTGGGACCT
 TGTGCATGCATGGTCCTCTCTGAGCCTTGGTTTCTTCAGCAGTGAGATGCTCAGAATAACTG
 CTGTCTCCCATGATGGTGTGGTACAGCGAGCTGTTGTCTGGCTATGGCATGGCTGTGCCGGG
 GGTGTTTGCTGAGGGCTTCCTGTGCCAGAGCCCAGCCAGAGAGCAGGTGCAGGTGTCATCCC
 GAGTTCAGGCTCTGCACGGCATGGAGTGGGAACCCACACAGCTGCTGCTACAGGACCTGGGA
 TTGCCTGGGACTCCACCTTCCTATCAATTCTCATGGTAGTCCAAACTGCAGACTCTCAAAC
 TTGCTCATTT

MSRYLLPLSALGTVAGAAVLLKDYVTGGACPSKATIPGKTVIVTGANTGIGKQTALELARRG
GNIILACRDM EKCEAAAKDIRGETLNHHVNARHLDLASLKSIREFAAKIIEEEERV DILINN
AGVMRCPHWTTEDGFEMQFGVNHLGHFLLTNLLLDKLKASAPSRIINLSSLAHVAGHIDFDD
LNWQTRKYNTKAAYCQSKLAIVLFTKELSRRLQSGSVTVNALHPGVARTELGRHTGIHGSTF
SSTTLGPIFWLLVKSPELAAQPSTYLAVAEELADVSGKYFDGLKQKAPAPEAEDEEVARRLW
AESARLVGLEAPSVREQPLPR

amino acids 1-17

FIGURE 67

GAAGTTTCGCGAGCGCTGGC**ATG**TGGTCCTGGGGCGCGGCTGGCGGCGCTGCTGGCGGTGCTG
 GCGCTCGGGACAGGAGACCCAGAAAGGGCTGCGGCTCGGGGCGACACGTTCTCGGCGCTGAC
 CAGCGTGGCGCGCGCCCTGGCGCCCGAGCGCCGGCTGCTGGGGCTGCTGAGGCGGTACCTGC
 GCGGGGAGGAGGCGCGGCTGCGGGACCTGACTAGATTCTACGACAAGGTACTTTCTTTGTCAT
 GAGGATTCACAACCCCTGTGGCTAACCCCTCTGCTTGCATTTACTCTCATCAAACGCCTGCA
 GTCTGACTGGAGGAATGTGGTACATAGTCTGGAGGCCAGTGAGAACATCCGAGCTCTGAAGG
 ATGGCTATGAGAAGGTGGAGCAAGACCTTCCAGCCTTTGAGGACCTTGAGGGAGCAGCAAGG
 GCCCTGATGCGGCTGCAGGACGTGTACATGCTCAATGTGAAAGGCCTGGCCCGAGGTGTCTT
 TCAGAGAGTCACTGGCTCTGCCATCACTGACCTGTACAGCCCCAAACGGCTCTTTTCTCTCA
 CAGGGGATGACTGCTTCCAAGTTGGCAAGGTGGCCTATGACATGGGGGATTATTACCATGCC
 ATTCCATGGCTGGAGGAGGCTGTCACTCTTCCGAGGATCTTACGGAGAGTGGAAGACAGA
 GGATGAGGCAAGTCTAGAAGATGCCTTGGATCACTTGGCCTTTGCTTATTTCCGGGCAGGAA
 ATGTTTCGTGTGCCCTCAGCCTCTCTCGGGAGTTTCTTCTCTACAGCCCAGATAATAAGAGG
 ATGGCCAGGAATGTCTTGAATATGAAAGGCTCTTGGCAGAGAGCCCCAACCCACGTGGTAGG
 TGAGGCTGTATCCAGAGGCCCAATATACCCACCTGCAGACCAGAGACACCTACGAGGGGC
 TATGTGAGACCCTGGGTTCAGCCCACTCTCTACCAGATCCCTAGCCTCTACTGTTCTCTAT
 GAGACCAATTCCAACGCCTACCTGCTGCTCCAGCCCATCCGGAAGGAGGTCTCCACCTGGA
 GCCCTACATTGCTCTCTACCATGACTTCGTCACTGACTCAGAGGCTCAGAAAATTAGAGAAC
 TTGCAGAACCATGGCTACAGAGGTCACTGGTGGCATCAGGGGAGAAGCAGTTACAAGTGGAG
 TACCGCATCAGCAAAAGTGCCTGGCTGAAGGACACTGTTGACCCAAAACCTGGTGACCCCTCAA
 CCACCGCATTGCTGCCCTCACAGGCCTTGATGTCCGGCCTCCCTATGCAGAGTATCTGCAGG
 TGGTGAACATATGGCATCGGAGGACACTATGAGCCTCACTTTGACCATGCTACGTCACCAAGC
 AGCCCCCTCTACAGAATGAAGTCAGGAAACCGAGTTGCAACATTTATGATCTATCTGAGCTC
 GGTGGAAGCTGGAGGAGCCACAGCCTTCATCTATGCCAACCTCAGCGTGCCTGTGGTTAGGA
 ATGCAGCACTGTTTTGGTGGAACTGCACAGGAGTGGTGAAGGGGACAGTGACACACTTCAT
 GCTGGCTGTCTGTCTGGTGGGAGATAAGTGGGTGGCCAACAAGTGGATACATGAGTATGG
 ACAGGAATTCGCGAGACCCTGCAGCTCCAGCCCTGAAGACT**TGA**ACTGTTGGCAGAGAGAAGC
 TGGTGGAGTCCTGTGGCTTTCCAGAGAAGCCAGGAGCCAAAAGCTGGGGTAGGAGAGGAGAA
 CAGCAGAGCAGCCTCCTGGAAGAAGGCCTTGTCAGCTTTGTCTGTGCCTCGCAAATCAGAGGC
 AAGGGGAGAGGTTGTTACCAGGGGACACTGAGAATGTACATTTGATCTGCCCCAGCCACGGAA
 GTCAGAGTAGGATGCACAGTACAAAGGAGGGGGGAGTGGAGGCCTGAGAGGGGAAGTTTCTGG
 AGTTCAGATACTCTCTGTTGGGAACAGGACATCTCAACAGTCTCAGGTTTCGATCAGTGGGTC
 TTTTGGCACTTTGAACCTTGACCACAGGGACCAAGAAGTGGCAATGAGGACACCTGCAGGAG
 GGGCTAGCCTGACTCCCAGAACTTTAAGACTTTCTCCCCACTGCCTTCTGCTGCAGCCCAAG
 CAGGGAGTGTCCCCCTCCCAGAACATATCCCAGATGAGTGGTACATTATATAAGGATTTT
 TTTAAGTTGAAAACAACCTTTCTTTTCTTTTGTATGATGGTTTTTTAACACAGTCATTAAAA
 ATGTTTATAAATCAAAA

FIGURE 68

MGPGARLAALLAVLALGTGDPERAAARGDTFSALTSVARALAPERRLLGLLRRYLRGEEARL
RDLTRFYDKVLSLHEDSTTPVANPLLAFTLIKRLQSDWRNVVHSLEASENIRALKDGYEKVE
QDLPAFEDLEGAARALMRLQDVYMLNVKGLARGVFQRVGTGSAITDLYSPKRLFSLTGDDCFQ
VGKVAYDMGDYYHAIPWLEEAVSLFRGSYGEWKTEDEASLEDALDHLAFAYFRAGNVSCALS
LSREFLLYSPDNKRMARNVLKYERLLAESPNHVVAEAVIQRPNI PHLQTRDTYEGLCQTLGS
QPTLYQIPSLYCSYETNSNAYLLLQPIRKEVIHLEPYIALYHDFVSDSEAQKIRELAEPWLQ
RSVVASGEKQLQVEYRISKSAWLKDTVDPKLVTLNHRIAALTGLDVRPPYAEYLQVVNYGIG
GHYEPHFDHATSPSSPLYRMKSGNRVATFMIYLSSVEAGGATAFIYANLSVPVVRNAALFWW
NLHRSGEGDSDTLHAGCPVLVGDKWVANKWIHEYGQEFRRPCSSSPED

Signal peptide:

amino acids 1-19

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 69

GAGATAGGGAGTCTGGGTTTAAAGTTCCTGCTCCATCTCAGGAGCCCCCTGCTCCCACCCCTAG
 GAAGCCACCAGACTCCACGGTGTGGGGCCAATCAGGTGGAATCGGCCCTGGCAGGTGGGGCC
 ACGAGCGCTGGCTGAGGGACCGAGCCGGAGAGCCCCGGAGCCCCCGTAACCCGCGCGGGGAG
 CGCCCAGGATGCCGCGCGGGGACTCGGAGCAGGTGCGCTACTGCGCGCGCTTCTCCTACCTC
 TGGCTCAAGTTTTTCACTTATCATCTATTCCACCGTGTTCTGGCTGATTGGGGCCCTGGTCCT
 GTCTGTGGGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAAACCCTTGAAAGTGCCTTCC
 TGGCTCCAGCCATCATCCTCATCCTCCTGGGCGTCGTCATGTTTCATGGTCTCCTTCATTGGT
 GTGCTGGCGTCCCTCCGTGACAACCTGTACCTTCTCCAAGCATTTCATGTACATCCTTGGGAT
 CTGCCTCATCATGGAGCTCATTGGTGGCGTGGTGGCCTTGACCTTCCGGAACCAGACCATTG
 ACTTCCTGAACGACAACATTTCGAAGAGGAATTGAGAACTACTATGATGATCTGGACTTCAAA
 AACATCATGGACTTTGTTTCAGAAAAAGTTCAAGTGCTGTGGCGGGGAGGACTACCGAGATTG
 GAGCAAGAATCAGTACCACGACTGCAGTGCCCCCTGGACCCCTGGCCTGTGGGGTGCCCTACA
 CCTGCTGCATCAGGAACACGACAGAAGTTGTCAACACCATGTGTGGCTACAAAACCTATCGAC
 AAGGAGCGTTTTCAGTGTGCAGGATGTCTACGTGCGGGGCTGCACCAACGCCGTGATCAT
 CTGGTTTCATGGACAACCTACACCATCATGGCGTGCATCCTCCTGGGCATCCTGCTTCCCCAGT
 TCCTGGGGGTGCTGCTGACGCTGCTGTACATCACCCGGGTGGAGGACATCATCATGGAGCAC
 TCTGTCACTGATGGGCTCCTGGGGCCCCGGTGCCAAGCCAGCGTGGAGGGCGGCAGGCACGGG
 ATGCTGCTTGTGCTACCCCAATTAGGGGCCCAGCCTGCCATGGCAGCTCCAACAAGGACCGTC
 TGGGATAGCACCTCTCAGTCAACATCGTGGGGCTGGACAGGGCTGCGGCCCCCTCTGCCACA
 CTCAGTACTGACCAAAGCCAGGGCTGTGTGTGCCTGTGTGTAGGTCCCACGGCCTCTGCCTC
 CCCAGGGAGCAGAGCCTGGGCCTCCCCTAAGAGGCTTTCCCCGAGGCAGCTCTGGAATCTGT
 GCCACCTGGGGCCTGGGGAACAAGGCCCTCCTTTCTCCAGGCCTGGGCTACAGGGGAGGGA
 GAGCCTGAGGCTCTGCTCAGGGCCCATTTTCATCTCTGGCAGTGCCTTGGCGGTGGTATTCAA
 GGCAGTTTTGTAGCACCTGTAATTGGGGAGAGGGAGTGTGCCCCTCGGGGCAGGAGGGAAGG
 GCATCTGGGGAAGGGCAGGAGGGAAGAGCTGTCCATGCAGCCACGCCCATGGCCAGGTGGC
 CTCTTCTCAGCCTCCCAGGTGCCTTGAGCCCTCTTGCAAGGGCGGCTGCTTCCTTGAGCCTA
 GTTTTTTTTTTACGTGATTTTTGTAAACATTCATTTTTTTGTACAGATAACAGGAGTTTCTGAC
 TAATCAAAGCTGGTATTTCCCCGCATGTCTTATTCTTGCCCTTCCCCCAACCAGTTTGTAA
 TCAAACAATAAAAACATGTTTTGTTTTGTTTTTAAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64863

><subunit 1 of 1, 294 aa, 1 stop

><MW: 33211, pI: 5.35, NX(S/T): 3

MPRGDSEQVRYCARFSYLWLKFSLLIYSTVFWLIGALVLSVGIYAEVERQKYKTLESAFLAP
AIILILLGVVMFMVSFIGVLASLRDNLYLLQAFMYILGICLIMELIGGVVALTFRNQTIDFL
NDNIRRGIEENYYDDLDFKNIMDFVQKKFKCCGGEDYRDWSKNQYHDCSAPGPLACGVPYTCC
IRNTTEVVNTMCGYKTIDKERFSVQDVIIYVRGCTNAVIIWFMDNYTIMACILLGILLPQFLG
VLLTLLYITRVEDIIMEHSVTDGLLGPGAKPSVEAAGTGCCLCYPN

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 22-42, 57-85, 93-116, 230-257

FIGURE 71

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGA
 CACCTGGGAAGATGGCCGGCCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACC
 TTGATCCAAGCCACCCTCAGTCCCAGTTCATCCTCGGCCCAAAAGTCATCAAAGA
 AAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGC
 TCAGTGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACC
 GTCCTGAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAA
 GCCCTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCA
 ACACGCCCCCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATC
 CGCATGGACACCAGTGCAAGTGGCCCCACCCGCCTGGTCTCAGTGACTGTGCCACCAGCCA
 TGGGAGCCTGCGCATCCAAGTGTGTATAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGC
 AGGTTCATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTG
 ATCGAGGCTTCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCATTTTC
 CCTCAGCATTTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTTC
 AGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAAC
 TCTGCAGCTTCCCTGACAATGCCACCCCTGGACAACATCCCGTTCAGCCTCATCGTGAGTCA
 GGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGAATTCATGGTCTCTGTTGG
 ACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAG
 GCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTT
 TTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTTCCCTCCA
 GTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTAC
 ACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGAT
 GAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCC
 ACTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTG
 AAGGCCTTGGGATTTCGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCC
 AGCCTCCTTGTGGAAACCCAGCTCTCCTGTCTCCCAGTGAAGACTTGGATGGCAGCCATCAG
 GGAAGGCTGGGTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAAT
 CAATAAACACTTGCCTGTGAAAAA

FIGURE 72

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881

><subunit 1 of 1, 484 aa, 1 stop

><MW: 52468, pI: 7.14, NX(S/T): 3

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQPLLSAM
 REKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPL
 VKTIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMN
 LLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITIQLYL
 GAKLLDSQGKVTKWFNNSAASLTMPITLDNIPFSLIVSQDVVKAABAVALSPEEFMVLLDSVL
 PESAHRLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEAL
 RPLFTLGIEASSEAQFYTKGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSIL
 LPNQNGKLRSGVPVSLVKALGFEEAESSLTKDALVLTTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 73

GAGCGAAC**ATGG**CAGCGCGTTGGCGGTTTTGGTGTGTCTCTGTGACCATGGTGGTGGCGCTG
 CTCATCGTTTGGCAGGTTCCCTCAGCCTCTGCCCAAAGAAAGAAGGAGATGGTGTTATCTGA
 AAAGGTTAGTCAGCTGATGGAATGGACTAACAAAAGACCTGTAATAAGAATGAATGGAGACA
 AGTTCCGTCGCCTTGTGAAAGCCCCACCGAGAAATTACTCCGTTATCGTCATGTTCACTGCT
 CTCCAACCTGCATAGACAGTGTGTCGTTTGAAGCAAGCTGATGAAGAATTCCAGATCCTGGC
 AACTCCTGGCGATACTCCAGTGCATTACCAACAGGATATTTTTTGGCATGGTGGATTTTG
 ATGAAGGCTCTGATGTATTTAGATGCTAAACATGAATTCAGCTCCAACTTTCATCAACTTT
 CCTGCAAAGGGAAACCCAAACGGGGTGATACATATGAGTTACAGGTGCGGGGTTTTTCAGC
 TGAGCAGATTGCCCGGTGGATCGCCGACAGAACTGATGTCAATATTAGAGTGATTAGACCCC
 CAAATTATGCTGGTCCCCTTATGTTGGGATTGCTTTTTGGCTGTTATTGGTGGACTTGTGTAT
 CTTTCGAAGAAGTAATATGGAATTTCTCTTTAATAAACTGGATGGGCTTTTGCAGCTTTGTG
 TTTTGTGCTTGCTATGACATCTGGTCAAATGTGGAACCATATAAGAGGACCACCATATGCC
 ATAAGAATCCCCACACGGGACATGTGAATTATATCCATGGAAGCAGTCAAGCCCAGTTTGT
 GCTGAAACACACATTGTTCTTCTGTTTAAATGGTGGAGTTACCTTAGGAATGGTGCTTTTATG
 TGAAGCTGCTACCTCTGACATGGATATTGGAAGCGAAAGATAATGTGTGTGGCTGGTATTG
 GACTTGTTGTATTATTCTTCAGTTGGATGCTCTCTATTTTTTAGATCTAAATATCATGGCTAC
 CCATACAGCTTTCTGATGAGT**TAA**AAAGGTCCCAGAGATATATAGACACTGGAGTACTGGAA
 ATTGAAAAACGAAAATCGTGTGTGTTTGAAAAGAAGAATGCAACTTGTATATTTTGTATTAC
 CTCTTTTTTTCAAGTGATTTAAATAGTTAATCATTTAACCAAAGAAGATGTGTAGTGCCTTA
 ACAAGCAATCCTCTGTCAAATCTGAGGTATTTGAAAATAATTATCCTCTTAACCTTCTCTT
 CCCAGTGAACCTTTATGGAACATTTAATTTAGTACAATTAAGTATATTATAAAAAATTGTAAAA
 CTACTACTTTGTTTTAGTTAGAACAAAGCTCAAACTACTTTAGTTAACTTGGTCATCTGAT
 TTTATATTGCCTTATCCAAAGATGGGGAAAGTAAGTCCTGACCAGGTGTTCCACATATGCC
 TGTTACAGATAACTACATTAGGAATTCATTCTTAGCTTCTTCATCTTTGTGTGGATGTGTAT
 ACTTTACGCATCTTTCCCTTTTGAGTAGAGAAATTATGTGTGTGTCATGTGGTCTTCTGAAAATG
 GAACACCATTCTTCAGAGCACACGTCTAGCCCTCAGCAAGACAGTTGTTTCTCCTCCTCCTT
 GCATATTTCCCTACTGCGCTCCAGCCTGAGTGATAGAGTGAGACTCTGTCTCAAAAAAAGTA
 TCTCTAAATACAGGATTATAATTTCTGCTTGAGTATGGTGTAACTACCTTGTATTTAGAAA
 GATTTTCAGATTCATTCCATCTCCTTAGTTTTCTTTTAAGGTGACCCATCTGTGATAAAAAATA
 TAGCTTAGTGCTAAAAATCAGTGTAACCTTATACATGGCCTAAAAATGTTTCTACAAATTAGAGT
 TTGTCACTTATTCATTTGTACCTAAGAGAAAAATAGGCTCAGTTAGAAAAGGACTCCCTGG
 CCAGGCGCAGTGACTTACGCCTGTAATCTCAGCACTTTGGGAGGCCAAGGCAGGCAGATCAC
 GAGGTGAGGAGTTTCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATAT
 AAAAATTAGCTGGGTGTGGTGGCAGGAGCCTGTAATCCCAGCTACACAGGAGGCTGAGGCAC
 GAGAATCACTTGAACCTCAGGAGATGGAGGTTTCAGTGAGCCGAGATCACGCCACTGCACTCC
 AGCCTGGCAACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAA

FIGURE 74

MAARWRFWCVSVTMVALLIVCDVPSASAQRKKEMVLSEKVSQLMWETNKRVPVIRMNGDKFR
RLVKAPPRNYSVIVMFTALQLHRQCVVCKQADEEFQILANSWRYSSAFTNRIFAMVDFDEG
SDVFQMLNMNSAPTFINFPAGKPKRGDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNY
AGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFALCFVLAMTSGQMWNHIRGPPYAHKN
PHTGHVNYIHGSSQAQFVAETHIVLLFNGGVTLGMVLLCEAATSDMDIGKRKIMCVAGIGLV
VLFFSWMLSIFRSKYHGYPSFLMS

Signal peptide:

amino acids 1-29

Transmembrane domains:

amino acids 183-205, 217-237, 217-287, 301-321

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 75

[illegible]

FIGURE 76

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64885

<subunit 1 of 1, 536 aa, 1 stop

<MW: 61450, pI: 9.17, NX(S/T): 7

MLLLWVSVVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDFSFDGRLTFHPGSQVVKLPFINF
MKTRGTSFLNAYTNSPICCPSPRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGYRTQ
KFGKLDYTSGHHSISNRVEAWTRDVAFLLRQEGRPMVNLIRNRTKVRVMERDWQNTDKAVNW
LRKEAINYTEPFVIYLG LNLPHYPSPSSGENFGSSTFHTSLYWLEKVS HDAIKIPKWSPLS
EMHPVDYYSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQLDLLQKTIVIIY
SSDHGELAMEHRQFYKMSMYEASAHVPLLMMGPGIKAGLQVSNVSLVDIYPTMLDIAGIPL
PQNLSGYSLLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYS DGAS
ILPQLFDLSSDPDEL TNVAVKFPEITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIG
QNYSNVIANLRWHQDWQKEPRKYENAIDQWLKTHMNPRAV

Important features:**Signal peptide:**

amino acids 1-15

N-glycosylation sites.

amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,
498-501

Sulfatases proteins:

amino acids 286-315, 359-369, 78-97

FIGURE 77

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAG
 AGCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATG**
 GCCTCTCTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACT
 GGTGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAG
 CAGTTGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAG
 TGTGACATCTATAGCACCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGAT
 GGTGACATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCA
 CAGTCTTCTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTC
 ATCCTTGGAGGCCTCCTGGGATTCATTCTCTGTTGCCTGGAATCTTCATGGGATCCTACGGGA
 CTTCTACTCACCACTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGG
 GCATTATTTCTTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCTCTGCTCATCC
 CAGAGAAATCGCTCCAACCTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTC
 TCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGT
 ATGTG**TGA**AAGAACCAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAG
 CACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGTGTGAGAAGGTGCTGCTGAGG
 ATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATG
 CAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTC
 CCCTGCCCTAAGTCCCCAACCTCAACTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGG
 ATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCCTG
 ACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGG
 GGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCC
 TCCAAAGAAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTTATGACTCCACAGTGTCC
 AGACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAG
 GATGCAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 78

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64886
><subunit 1 of 1, 230 aa, 1 stop
><MW: 24549, pI: 8.56, NX(S/T): 1
MASLGLQLVGYYLGLLGLLGTLLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGIT
QCDIYSTLLGLPADIQAAQAMMVTSSAISLACIISVVGMRCTVFCQESRAKDRVAVAGGVF
FILGGLLGFIPIVAVNHLGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCS
SQNRNSNYDAYQAQPLATRSSPRPGQPPKVKSEFNSSYSLTGYV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

GCACCTGCTGCTGTCCCATCAGCTGCTCTGAAGCTCCATGGTGCCCGAGAATCTTCGCTCCTGC
TTATGTGTCTAGTCTGTCTCCTCCTCTTGTGTCCAAGGGAAGTCATCGCTCCCGCTGGCTCAG
AACCATGGCTGTGCCAGCCGGCACCCAGGTGTGGAGACAAGATCTACAACCCCTTGGAGCAG
TGCTGTTACAATGACGCCATCGTGTCCCTGAGCGAGACCCGCCAATGTGGTCCCCCCTGCAC
CTTCTGGCCCTGCTTTGAGCTCTGCTGTCTTGATTCTTTGGCCTCACAAACGATTTTGTTG
TGAAGCTGAAGGTTTCAAGGTGTGAATTCCCAGTGCCACTCATCTCCCATCTCCAGTAAATGT
GAAAGCAGAAGACGTTTTCCCTTGAGAAAGACATAGAAAGAAAATCAACTTTCCTAAGGCATC
TCAGAAACATAGGCTAAGGTAATATGTGTACCAAGTAGAGAAGCCTGAGGAATTTACAAAATG
ATGCAGCTCCAAGCCATTGTATGGCCCATGTGGGAGACTGATGGGACATGGAGAATGACAGT
AGATTATCAGGAAATAAATAAAGTGGTTTTTCCAATGTACACACCTGTAAAA

FIGURE 80

MVPRI FAPAYVSVCLLLLCPREVIAPAGSEPWLCQPAPRCGDKIYNPLEQCCYNDAIVSLSE
TRQCGPPCTFWPCFELCCLDSFGLTNDVFVVKLVQGVNSQCHSSPISSKCESRRRFP

Signal peptide:

amino acids 1-25

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 81

CTCCACTGCAACCACCCAGAGCC**ATG**GCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATT
TTCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCT
GATGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCCTGCAGCACTGTT
GCTATGATGATGCCGTCGTGCCCTTGGCCAGGACCCAGACGTGTGGAAACTGCACCTTCAGA
GTCTGCTTTGAGCAGTGCTGCCCCCTGGACCTTCATGGTGAAGCTGATAAACCAGAACTGCCGA
CTCAGCCCCGGACCTCGGATGACAGGCTTTGTGCGAGTGTGAG**TAA**TGGAACATCAGGGGAA
CGATGACTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTTACCTGAGA
TCTGGGATGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCCACTTCATT
CTGTGACCTGTCTGAGGCCCCACCCTGCAGCTGCCCTGAGGAGGCCCCACAGGTCCCCTTCTAG
AATTCTGGACAGCATGAGATGCGTGTGCTGATGGGGGCCAGGGACTCTGAACCCTCCTGAT
GACCCCTATGGCCAACATCAACCCGGCACCACCCCAAGGCTGGCTGGGGAACCCTTCACCCT
TCTGTGAGATTTTCCATCATCTCAAGTTCTCTTCTATCCAGGAGCAAAGCACAGGATCATAA
TAAATTTATGTACTTTATAAATGAAAA

82/249

FIGURE 82

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVP
LARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCD SARTSDDRLCRSVS

Signal peptide:

amino acids 1-24

100172621001

GGGGGCGGGGTGCCTGGAGCACGGGCGCTGGGGGCCCGCCGCGAGCGCTCACTCGCTCGCACTCAG
TCGCGGGAGGCTTCCCCGCGCCGGCCGCGTCCCGCCCGCTCCCCGGCACCAGAAGTTCCTCT
GCGCGTCCGACGGCGAC**CATG**GGGCGTCCCCACGGCCCTGGAGGCCGGCAGCTGGCGCTGGGGA
TCCCTGCTCTTCGCTCTCTTCCTGGCTGCGTCCCTAGGTCCGGTGGCAGCCTTCAAGGTCGC
CACGCCGTATTCCCTGTATGTCTGTCCCGAGGGGGCAGAACGTCACCCTCACCTGCAGGCTCT
TGGGCCCTGTGGACAAAGGGCACGATGTGACCTTCTACAAGACGTGGTACCGCAGCTCGAGG
GGCGAGGTGCAGACCTGCTCAGAGCGCCGGGCCATCCGCAACCTCACGTTCCAGGACCTTCA
CCTGCACCATGGAGGCCACCAGGCTGCCAACACCAGCCACGACCTGGCTCAGCGCCACGGGC
TGGAGTCGGCCTCCGACCACCATGGCAACTTCTCCATCACCATGCGCAACCTGACCCTGCTG
GATAGCGGCCTCTACTGCTGCCTGGTGGTGGAGATCAGGCACCACCACTCGGAGCACAGGGT
CCATGGTGCCATGGAGCTGCAGGTGCAGACAGGCCAAAGATGCACCATCCAACCTGTGTGGTGT
ACCCATCCTCCTCCCAGGATAGTGAAAACATCACGGCTGCAGCCCTGGCTACGGGTGCCTGC
ATCGTAGGAATCCTCTGCCTCCCCCTCATCCTGCTCCTGGTCTACAAGCAAAGGCAGGCAGC
CTCCAACCGCCGTGCCCAGGAGCTGGTGC GGATGGACAGCAACATTCAAGGGATTGAAAACC
CCGGCTTTGAAGCCTCACACCTGCCCAGGGGATACCCGAGGCCAAAGTCAGGCACCCCTG
TCCTATGTGGCCAGCGGCAGCCTTCTGAGTCTGGGCGGCATCTGCTTTTCGGAGCCCAGCAC
CCCCCTGTCTCCTCCAGGCCCGGAGACGTCTTCTTCCCATCCCTGGACCCTGTCCCTGACT
CTCCAAACTTTGAGGTCATC**TAG**CCCCAGCTGGGGGACAGTGGGCTGTTGTGGCTGGGTCTGG
GGCAGGTGCATTTGAGCCAGGGCTGGCTCTGTGAGTGGCCTCCTTGGCCTCGGCCCTGGTTC
CCTCCCTCCTGCTCTGGGCTCAGATACTGTGACATCCAGAAAGCCCAGCCCCTCAACCCCTC
TGGATGCTACATGGGGATGCTGGACGGCTCAGCCCCTGTTCCAAGGATTTTGGGGTGCTGAG
ATTCTCCCCTAGAGACCTGAAATTCACCAGCTACAGATGCCAAATGACTTACATCTTAAGAA
GTCTCAGAACGTCCAGCCCTTCAGCAGCTCTCGTTCTGAGACATGAGCCTTGGGATGTGGCA
GCATCAGTGGGACAAGATGGACACTGGGCCACCCCTCCAGGCACCAGACACAGGGCACGGTG
GAGAGACTTCTCCCCCGTGGCCGCCTTGGCTCCCCCGTTTTGCCCGAGGCTGCTCTTCTGTC
AGACTTCCTCTTTGTACCACAGTGGCTCTGGGGCCAGGCCTGCCTGCCCCACTGGCCATCGCC
ACCTTCCCCAGCTGCCTCCTACCAGCAGTTTCTCTGAAGATCTGTCAACAGGTTAAGTCAAT
CTGGGGCTTCCACTGCCTGCATTCCAGTCCCCAGAGCTTGGTGGTCCCGAAACGGGAAGTAC
ATATTGGGGCATGGTGGCCTCCGTGAGCAAATGGTGTCTTGGGCAATCTGAGGCCAGGACAG
ATGTTGCCCCACCCACTGGAGATGGTGTGAGGGAGGTGGGTGGGGCCTTCTGGGAAGGTGA
GTGGAGAGGGGCACCTGCCCCCGCCCTCCCCATCCCCTACTCCCACTGCTCAGCGCGGGCC
ATTGCAAGGGTGCCACACAATGTCTTGTCCACCCTGGGACACTTCTGAGTATGAAGCGGGAT
GCTATTAATAAACTACATGGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

FIGURE 84

```
></usr/segdb2/sst/DNA/Dnaseqs.min/ss.DNA64897
```

```
><subunit 1 of 1, 311 aa, 1 stop
```

><MW: 33908, pI: 6.87, NX(S/T): 6

MGVPTALEAGSWRWGSLFLFALFLAASLGPVAAFVKVATPYSLYVCPEGQNVTLTCRLLGPVDK
GHDVTFYKTYWRSSRGEVQTCSERRPIRNLTFQDLHLHHGGHQAANTSHDLAQRHGLEASD
HHGNFSITMRNLTLTLDGLYCCLVVEIRHHHSEHRVHGAMELQVQTGKDAPSNVVPSSSQ
DSENITAAALATGACIVGILCLPLILLLVYKQRQAASNRRAEQELVRMDSNIQGIENPGFEAS
PPAQGIPEAKVRHPLSYVAQRQPSESGRHLLSEPSTPLSPPGPDVFFPSLDPVPDSPNFEVI

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 190-216

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCT
TTCCCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCTCC
TTTCCTGCCCCACGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGG
GGTCTGTGGGTTGATCTGTGGCCCCTGTGCCTCCGTGTCTTTTCGTCTCCCTTCCTCCCGA
CTCCGCTCCCGGACCAGCGGCCTGACCCTGGGGAAAGG**ATG**GTTCCCGAGGTGAGGGTCCTC
TCCTCCTTGCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCC
AGACATGTTCTGCCTTTTCCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACT
TGGAGCCACAAGGCCTGATGTACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGT
TGTTACCGCCTCCACTGTCCGCCTGTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATG
CTGTCCCAAGTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGGCCCCACCAAAGTCCTGCC
AGCACAACGGGACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCCTCC
CGCCTGCCCAACCAGTGTGTCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCTCAC
AACCTGCCCCGAACCAGGCTGCCCAGCACCCCTCCCCTGCCAGACTCCTGCTGCCAAGCCT
GCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTG
AGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAGAGAGGCCCGGGCACCCACAGC
CCCCACTGGCCTCAGCGCCCCTCTGAGCTTCATCCCTCGCCACTTCAGACCCAAGGGAGCAG
GCAGCACAACCTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGGCGGG
AAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCCTTGCCCTG
CATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGT
ACCCCTGCCGTCACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCCAGAGGACAAA
GCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTCTCT
CGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGG
CCTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAG
AGAGGTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGA
AAGTCAGGAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCAC
GAAGGTCACTGGAACGTCTTCCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGA
CAAAGTGACCAAGACATAACAAAGACCT**TAA**CAGTTGCAGATATGAGCTGTATAATTGTTGTT
ATTATATATTAATAAATAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 86

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64902
```

```
><subunit 1 of 1, 451 aa, 1 stop
```

><MW: 49675, pI: 7.15, NX(S/T): 1

MVPEVRVLSSLLGLALLWFLPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCT
 CSEGAHVSCYRLHCPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIF
 SAHELFP SRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEED
 SVQSLHGVRHPQDPCSSDAGRKRGP GTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKH
 KKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHPKEKVAGKC
 CKICPEDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRALEHEASDLVEIYLWKLV
 KDEETEAQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDPG
 AEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

[illegible]

FIGURE 88

MDSLRLKMLISVAMLGAGAGVGYALLVIVTPGERRRKQEMLKEMPLQDPRSREEAARTQQLLLA
TLQEAATTQENVAWRKNWMVGEGGASGRSP

Signal peptide:

amino acids 1-18

Downloaded from www.ashg.org

FIGURE 89

[illegible]

FIGURE 90

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64905

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46038, pI: 6.50, NX(S/T): 2

MGPSTPLLILFLLSWGPLQGQQHHLVEYMERRLAAL EERLAQCQDQSSRHAAELRDFKNKM
LPLLEVAEKEREALRTEADTISGRVDRLEREVDYLETQNPALPCVEFDEKVTGGPGTKGKGR
RNEKYDMVTDCGYTISQVRSMKILKRFGGPAGLWTKDPLGQTEKIYVLDGTQNDTAFVFPRL
RDFTLAMAARKASRVRVFPFPWVG TGQLVYGGFLYFARRPPGRPGGGGEMENTLQLIKFHLAN
RTVVDSSVFPAEGLIPPYGLTADTYIDLVADEEGLWAVYATREDDRHLCLAKLDPQTL DTEQ
QWDTPCPRENAEAAAFVICGTLYVVYNTRPASRARIQCSFDASGTLTPERAALPYFPRRYGAH
ASLRYNPRERQLYAWDDGYQIVYKLEMRKKEEEV

Important features:**Signal peptide:**

amino acids 1-21

N-glycosylation sites.

amino acids 177-180, 248-251

1004554001

FIGURE 91

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAG
 AGCCTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCAT
 CCGTCACCTCTCCTGTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCC**ATGG**
 CTCTCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTT
 GGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAGGACGCAGCATTCTCCTGTTTCCTGTC
 TCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCCTCAGGGGCCAGTTCTCTAGCGTGG
 TCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGG
 ACAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTAC
 TGTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCA
 TCTGGGAGCTACAGGTGTCAGCACTGGGCTCAGTTCCTCTCATTTCATCACGGGATATGTT
 GATAGAGACATCCAGCTACTCTGTCAGTCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTG
 GAAAGGTCCACAAGGACAGGATTTGTCCACAGACTCCAGGACAAACAGAGACATGCATGGCC
 TGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAACGCCGGGAGCATATCCTGTTCCATG
 CGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTTTTCGA
 GCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTATTTTTTG
 GCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCCAGTGGAATCCAGGCGGAACTGGAC
 TGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGAC
 TCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCCATA
 GAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCT
 TCTCAGAGTTTCCAAGCAGGGAAACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTG
 GCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCCC
 ATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTACATTAAATCCCCGT
 TTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCCTGGACTATGAGTG
 TGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGT
 TTGAAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAACCTCC
 ATAGTCATCTGCCCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGC
 AATCCCAGAGACAAGCAACAGTGAGTCCTCCTCACAGGCAACCACGCCCTTCCTCCCCAGGG
 GTGAAATG**TAG**GATGAATCACATCCCACATTCTTCTTTAGGGATATTAAGGTCTCTCTCCA
 GATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCAGATGAAGGGGGACTGGCCTGTCC
 ACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGGCTGACATTACATTT
 AGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAGAACCG
 TCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGC
 TTAGATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTA
 AAAAAA

FIGURE 92

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCLSPKTNAEAMEVRFFRGQFSS
VVHLYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQK
AIWELQVSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGQDLSTDSRTNRDMH
GLFDVEISLTVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLF
FGIVGLKIFFSKFQWKIQAELDWRRKHGQAEIRDARKHAVEVTLPETAHPKLCVSDLKTVT
HRKAPQEVPHSEKRFTRKSVVASQS FQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLS
PDHGYWVLRNLNGEHL YFTLNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTC
RFEGLLRPYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAI PETSNSESSSQATTPFLP
RGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

CGCATGGTGCGCCCGGTGGCGGTGGCGCGCGGTTGCGGAGGCTTCCTTGGTTCGGATTGCA
ACGAGGAGAAGATGACTGACCAACCGACTGGCTGAATGAATGAATGGCGGAGCCGAGCGCGC
CATGAGGAGCCTGCCGAGCCTGGGCGGCCTCGCCCTGTTGTGCTGCGCCGCCGCCGCCGCCG
CCGTCGCCTCAGCCGCCTCGGCGGGGAATGTCACCGGTGGCGGCGGGGCCGCCGGGGCAGGTG
GACGCGTCGCCGGGGCCCCGGGTTGCGGGGCGAGCCCAGCCACCCCTTCCCTAGGGCGACGGC
TCCCACGGCCCAGGCCCCGAGGACCGGGCCCCCGCGCGCCACCGTCCACCGACCCCTGGCTG
CGACTTCTCCAGCCAGTCCCCGGAGACCACCCCTCTTTGGGCGACTGCTGGACCCCTCTTCC
ACCACCTTTCAGGCGCCGCTCGGCCCTCGCCGACCACCCCTCCGGCGGCGGAACGCACCTTC
GACCACCTCTCAGGCGCCGACCAGACCCGCGCCGACCACCCCTTTCGACGACCACTGGCCCCG
CGCCGACCACCCCTGTAGCGACCACCGTACCGGCGCCACGACTCCCCGGACCCCGACCCCC
GATCTCCCCAGCAGCAGCAACAGCAGCGTCTCCCCACCCACCTGCCACCGAGGCCCCCTC
TTCGCCTCCTCCAGAGTATGTATGTAACCTGCTCTGTGGTTGGAAGCCTGAATGTGAATCGCT
GCAACCAGACCACAGGGCAGTGTGAGTGTGCGCCAGGTTATCAGGGGCTTCAGTGTGAAACC
TGCAAAGAGGGCTTTTACCTAAATTACACTTCTGGGCTCTGTGAGCCATGTGACTGTAGTCC
ACATGGAGCTCTCAGCATACCGTGCAACAGG**TAA**GCAACAGAGGGTGGAAGTGAAGTTTATT
TTATTTTAGCAAGGGAAAAAAAAGGCTGCTACTCTCAAGGACCATACTGGTTTAAACAAAG
GAGGATGAGGGTCATAGATTTACAAAATATTTTATATACTTTTATTCTCTTACTTTTATATGT
TATATTTAATGTGAGGATTTAAAAACATCTAATTTACTGATTTAGTTCTTCAAAAGCACTAG
AGTCGCCAATTTTTCTCTGGGATAATTTCTGTAAATTTTCATGGGAAAAAATTATTGAAGAAT
AAATCTGCTTTCTGGAAGGGCTTTCAGGCATGAAACCTGCTAGGAGGTTTAGAAATGTTCTT
ATGTTTATTAATATACCATTTGGAGTTTGAGGAAATTTGTTGTTTGGTTTATTTTTCTCTCTA
ATCAAATTTCTACATTTGTTTTCTTTGGACATCTAAAGCTTAACCTGGGGGTACCCTAATTTA
TTTAACTAGTGGTAAGTAGACTGGTTTTACTCTATTTACCAGTACATTTTTTGAGACCAAAG
TAGATTAAGCAGGAATTATCTTTAAACTATTATGTTATTTGGAGGTAATTTAATCTAGTGGA
ATAATGTACTGTTATCTAAGCATTTGCCTTGTAAGTACTGCACTGAAAGTAATTATTCTTTGACCT
TATGTGAGGCACTTGGCTTTTTGTGGACCCCAAGTCAAAAAACTGAAGAGACAGTATTAAAT
AATGAAAAAATAATGACAGGTTATACTCAGTGTAACCTGGGTATAACCCAAGATCTGCTGC
CACTTACGAGCTGTGTTTCTTGGGCAAGTAATTTCTTTCACTGAGCTTGTTTCTTCTCAAG
GTTGTTGTGAAGATTAAATGAGTTGATATATATAAAATGCCTAGCACATGTCACTCAATAAA
TTCTGGTTTGTTTTAATTTCAAAGGAATATTATGGACTGAAATGAGAGAACATGTTTTAAGA
ACTTTTAGCTCCTTGACAAAGAAGTGCTTTTATACTTTAGCACTAAATATTTTAAATGCTTTA
TAAATGATATTATACTGTTATGGAATATTGTATCATATTGTAGTTTATTAAAAATGTAGAAG
AGGCTGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAGGCCAAGGCGGGTGGAT
CACTTGAGGCCAGGAGTTCTAGATGAGCCTGGCCAGCACAGTGAACCCCGTCTCTACTAAA
AATACAAACAAATTAGCTGGGCGTGGTGGCACACACCTGTAGTCCAGCTACTCGGGAGGCT
GAGGCAGGAGAATCGGTTGAACCCGGGAGGTGGAGGTTGCAGTGAGCTGAGATCGCGCCACT
GCACTCCAGCCTGGTGAGAGAGGGGAGACTCTGTCTTAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 94

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952
><subunit 1 of 1, 258 aa, 1 stop
><MW: 25716, pI: 8.13, NX(S/T): 5
MRSLSLGLGGLALLCCAAAAA AVASAASAGNVTGGGGAAGQVDASPGPGLRGEP SHPFPRATA
PTAQAPRTGPPRATVHRPLAATSPAQSPETTPLWATAGPSSTTFQAPLGPSPTTPPAAERTS
TTSQAPTRPAPTTLSTTTGPAPTTPVATTVPAPTTPRTPDLPSSSNSSVLPTPPATEAPS
SPPPEYVCNCSVVGSLNVNRCNQTTGQCECRPGYQGLHCETCKEGFYLNYSGLCQPCDCSP
HGALSIPCNR

Important features of the protein:**Signal peptide:**

amino acids 1-25

N-glycosylation sites.

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

EGF-like domain cysteine pattern signature.

amino acids 214-226.

TGCGGCGCAGTGTAGACCTGGGAGG**ATG**GGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTTGG
TCTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGG
CCCTGGTACGTGCTTGCGGTGGCCTCCCGGGAAAAGGGCTTTGCCATGGAGAAGGACATGAA
GAACGTCTGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCTGTCTCTC
AGCACGGGCTGGGAGGGTGTGACCAGAGTGTGATGGACCTGATAAAGCGAAACTCCGGATGG
GTGTTTGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCACCAACTTCAGAGA
CTATGCCATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAACACCGTGGAGCTGT
ACAGTCTGACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTACCAAGTGGAGCAGGAGC
CTGGGCTTCCTGTACAG**TAG**CAGGCCCAGCTGCAGAAGGACCTCACCTGTGCTCACAAAGAT
CCTTCTGTGAGTGCTGCGTCCCCAGTAGGGATGGCGCCACAGGGTCTGTGACCTCGGCCA
GTGTCCACCCACCTCGCTCAGCGGCTCCCGGGGCCAGCACCAGCTCAGAATAAAGCGATTC
CACAGCA

FIGURE 96

MGGLLLAFLALVSVPRQAVWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVT
LTPENNLRTLSSQHGLGGCDQSVMDLIKRNSGWVFENPSIGVLELWVLATNFRDYAIIFTQL
EFGDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

Signal peptide:

amino acids 1-20

protein sequence

FIGURE 97

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCCAGACATGCTGCTGCTGCTGCTGCC
 CCTGCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTT
 CCGTGACGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCAT
 GGCTGGATTTACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGA
 CCAGGATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACC
 GATTCCACCTCCTTGGGGACCCACATAACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGA
 AGAAGTGATGCGGGGAGATACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAA
 ACATCACCGGCTCTCTGTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAG
 GCACCCTGGAGTCCGGCTGCCCCCAGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAG
 GGGACACCCCCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCTGGACCCCTCCACCAC
 CCGCTCCTCGGTGCTCACCCCTCATCCCACAGCCCCAGGACCATGGCACCCAGCCTCACCTGTC
 AGGTGACCTTCCCTGGGGCCAGCGTGACCACGAACAAGACCGTCCATCTCAACGTGTCCTAC
 CCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGAGACGGCACAGTATCCACAGTCTTGGG
 AAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATG
 CAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGAGAGGCCTGACCCTGTGCCCC
 TCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGA
 ATTCACCTGCAGAGCTCAGAACCCCTCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGC
 AGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTTCGGGGGAGCTGGAGCCACAGCCCTG
 GTCTTCCTGTCCTTCTGCGTCATCTTCGTTGTAGTGAGGTCCTGCAGGAAGAAATCGGCAAG
 GCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTTCAGCCT
 CTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAGCT
 TCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGT
 GAAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCC
 ACAGATGAAGAACTGCAGAGACTCACCCCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGA
 GAAGTCAGAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACT
 ATGAATTATGTGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAAT
 CCACACTGTGCCCTCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIIYPGPVVHGYWF
REGANTDQDAPVATNNPARAVWEETRDRFHLGDPHTKNCTLSIRDARRSDAGRYFFRMEKG
SIKWNYPKHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVS
PLDPSTTRSSVLTLLIPQPQDHGTS LTCQVTFPGASVTTNKTVHLNVSYPPQNLTMTVFQGDG
TVSTVLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWV
HLRDAAEFTCRAQNPLGSQQVYLVNLSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRS
CRKKSARPAAGVGDGTGIEDANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGEDELQYA
SLSFQMVKPWDSRGQEATDTEYSEIKIHR

amino acids 1-15

amino acids 351-370

FIGURE 99

GACGCCCAGTGACCTGCCGAGGTCGGCAGCACAGAGCTCTGGAGATGAAGACCCTGTTTCCTG
GGTGTACGCTCGGCCTGGCCGCTGCCCTGTCCTTCACCCTGGAGGAGGAGGATATCACAGG
GACCTGGTACGTGAAGGCCATGGTGGTCGATAAGGACTTTCCGGAGGACAGGAGGCCCCAGGA
AGGTGTCCCCAGTGAAGGTGACAGCCCTGGGCGGTGGGAAGTTGGAAGCCACGTTCACCTTC
ATGAGGGAGGATCGGTGCATCCAGAAGAAAATCCTGATGCGGAAGACGGAGGAGCCTGGCAA
ATACAGCGCCTATGGGGGCAGGAAGCTCATGTACCTGCAGGAGCTGCCCAGGAGGGACCACT
ACATCTTTTACTGCAAAGACCAGCACCATGGGGGCCTGCTCCACATGGGAAAGCTTGTGGGT
AGGAATTCTGATACCAACCGGGAGGCCCTGGAAGAATTTAAGAAATTGGTGCAGCGCAAGGG
ACTCTCGGAGGAGGACATTTTCACGCCCCCTGCAGACGGGAAGCTGCGTTCCCGAACACTAGG
CAGCCCCCGGGTCTGCACCTCCAGAGCCCACCCTACCACCAGACACAGAGCCCGGACCACCT
GGACCTACCCTCCAGCCATGACCCTTCCCTGCTCCCACCCACCTGACTCCAAATAAAGTCCT
TTCCCCCAA

FIGURE 100

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65404

<subunit 1 of 1, 170 aa, 1 stop

<MW: 19457, pI: 9.10, NX(S/T): 0

MKTLFLGVTLGLAAALSFTLEEEDITGTWYVKAMVVDKDFPEDRRPRKVSPVKVTALGGGKL
EATFTFMREDRCIQKKILMRKTEEPGKYSAYGGRKLMYLQELPRRDHYIFYCKDQHHGGLLH
MGKLVGRNSDTNREALLEEFKKLVQRKGLSEEDIFTPLQTGSCVPEH

Important features:**Signal peptide:**

amino acids 1-17

100/249

FIGURE 101

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCAC
AGCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCTCCAGGCCATGAGGATTCTGCAGTTAA
TCCTGCTTGCTCTGGCAACAGGGCTTGTAAGGGGAGAGACCAGGATCATCAAGGGGTTCGAG
TGCAAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTGAGAAGACGCGGCTACTCTGTGG
GGCGACGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACA
TAGTTCACCTGGGGCAGCACAACTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCC
ACTGAGTCCTTCCCCACCCCGGCTTCAACAACAGCCTCCCCAACAAAGACCACCGCAATGA
CATCATGCTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCC
TCTCCTCACGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCC
AGCCCCCAGTTACGCCTGCCTCACACCTTGCGATGCGCCAACATCACCATCATTGAGCACCA
GAAGTGTGAGAACGCCTACCCCGGCAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGG
AAGGGGGCAAGGACTCCTGCCAGGGTGACTCCGGGGGGCCCTCTGGTCTGTAACCAGTCTCTT
CAAGGCATTATCTCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACAC
GAAAGTCTGCAAATATGTGGACTGGATCCAGGAGACGATGAAGAACAATTAGACTGGACCCA
CCCACCACAGCCCATCACCCCTCCATTTCCACTTGGTGTGTTGGTTCCTGTTCACTCTGTTAAT
AAGAAACCCTAAGCCAAGACCCTCTACGAACATTCTTTGGGCCTCCTGGACTACAGGAGATG
CTGTCACTTAATAATCAACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTT
GAAATATTGTGACTCTGGGAATGACAACACCTGGTTTGTCTCTGTTGTATCCCCAGCCCCA
AAGACAGCTCCTGGCCATATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 102

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65405

<subunit 1 of 1, 250 aa, 1 stop

<MW: 27466, pI: 8.87, NX(S/T): 4

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAH
 CLKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITW
 AVRPLTLSSRCVTAGTSCCLISGWGSTSSPQLRLPHTLRCANITIIIEHQKCENAYPGNITDTM
 VCASVQEGGKDSCQGDGGPLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDWIIQETMKNN

Important features:

Signal peptide:

amino acids 1-18

Serine proteases, trypsin family, histidine active site.

amino acids 58-63

N-glycosylation sites.

amino acids 99-102, 165-168, 181-184, 210-213

Glycosaminoglycan attachment site.

amino acids 145-148

Kringle domain proteins.

amino acids 197-209, 47-64

Serine proteases, trypsin family, histidine protein

amino acids 199-209, 47-63, 220-243

Apple domain proteins

amino acids 222-249, 189-222

1001
 1002
 1003
 1004
 1005
 1006
 1007
 1008
 1009
 1010
 1011
 1012
 1013
 1014
 1015
 1016
 1017
 1018
 1019
 1020
 1021
 1022
 1023
 1024
 1025
 1026
 1027
 1028
 1029
 1030
 1031
 1032
 1033
 1034
 1035
 1036
 1037
 1038
 1039
 1040
 1041
 1042
 1043
 1044
 1045
 1046
 1047
 1048
 1049
 1050
 1051
 1052
 1053
 1054
 1055
 1056
 1057
 1058
 1059
 1060
 1061
 1062
 1063
 1064
 1065
 1066
 1067
 1068
 1069
 1070
 1071
 1072
 1073
 1074
 1075
 1076
 1077
 1078
 1079
 1080
 1081
 1082
 1083
 1084
 1085
 1086
 1087
 1088
 1089
 1090
 1091
 1092
 1093
 1094
 1095
 1096
 1097
 1098
 1099
 1100
 1101
 1102
 1103
 1104
 1105
 1106
 1107
 1108
 1109
 1110
 1111
 1112
 1113
 1114
 1115
 1116
 1117
 1118
 1119
 1120
 1121
 1122
 1123
 1124
 1125
 1126
 1127
 1128
 1129
 1130
 1131
 1132
 1133
 1134
 1135
 1136
 1137
 1138
 1139
 1140
 1141
 1142
 1143
 1144
 1145
 1146
 1147
 1148
 1149
 1150
 1151
 1152
 1153
 1154
 1155
 1156
 1157
 1158
 1159
 1160
 1161
 1162
 1163
 1164
 1165
 1166
 1167
 1168
 1169
 1170
 1171
 1172
 1173
 1174
 1175
 1176
 1177
 1178
 1179
 1180
 1181
 1182
 1183
 1184
 1185
 1186
 1187
 1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449
 1450
 1451
 1452
 1453
 1454
 1455
 1456
 1457
 1458
 1459
 1460
 1461
 1462
 1463
 1464
 1465
 1466
 1467
 1468
 1469
 1470
 1471
 1472
 1473
 1474
 1475
 1476
 1477
 1478
 1479
 1480
 1481
 1482
 1483
 1484
 1485
 1486
 1487
 1488
 1489
 1490
 1491
 1492
 1493
 1494
 1495
 1496
 1497
 1498
 1499
 1500
 1501
 1502
 1503
 1504
 1505
 1506
 1507
 1508
 1509
 1510
 1511
 1512
 1513
 1514
 1515
 1516
 1517
 1518
 1519
 1520
 1521
 1522
 1523
 1524
 1525
 1526
 1527
 1528
 1529
 1530
 1531
 1532
 1533
 1534
 1535
 1536
 1537
 1538
 1539
 1540
 1541
 1542
 1543
 1544
 1545
 1546
 1547
 1548
 1549
 1550
 1551
 1552
 1553
 1554
 1555
 1556
 1557
 1558
 1559
 1560
 1561
 1562
 1563
 1564
 1565
 1566
 1567
 1568
 1569
 1570
 1571
 1572
 1573
 1574
 1575
 1576
 1577
 1578
 1579
 1580
 1581
 1582
 1583
 1584
 1585
 1586
 1587
 1588
 1589
 1590
 1591
 1592
 1593
 1594
 1595
 1596
 1597
 1598
 1599
 1600
 1601
 1602
 1603
 1604
 1605
 1606
 1607
 1608
 1609
 1610
 1611
 1612
 1613
 1614
 1615
 1616
 1617
 1618
 1619
 1620
 1621
 1622
 1623
 1624
 1625
 1626
 1627
 1628
 1629
 1630
 1631
 1632
 1633
 1634
 1635
 1636
 1637
 1638
 1639
 1640
 1641
 1642
 1643
 1644
 1645
 1646
 1647
 1648
 1649
 1650
 1651
 1652
 1653
 1654
 1655
 1656
 1657
 1658
 1659
 1660
 1661
 1662
 1663
 1664
 1665
 1666
 1667
 1668
 1669
 1670
 1671
 1672
 1673
 1674
 1675
 1676
 1677
 1678
 1679
 1680
 1681
 1682
 1683
 1684
 1685
 1686
 1687
 1688
 1689
 1690
 1691
 1692
 1693
 1694
 1695
 1696
 1697
 1698
 1699
 1700
 1701
 1702
 1703
 1704
 1705
 1706
 1707
 1708
 1709
 1710
 1711
 1712
 1713
 1714
 1715
 1716
 1717
 1718
 1719
 1720
 1721
 1722
 1723
 1724
 1725
 1726
 1727
 1728
 1729
 1730
 1731
 1732
 1733
 1734
 1735
 1736
 1737
 1738
 1739
 1740
 1741
 1742
 1743
 1744
 1745
 1746
 1747
 1748
 1749
 1750
 1751
 1752
 1753
 1754
 1755
 1756
 1757
 1758
 1759
 1760
 1761
 1762
 1763
 1764
 1765
 1766
 1767
 1768
 1769
 1770
 1771
 1772
 1773
 1774
 1775
 1776
 1777
 1778
 1779
 1780
 1781
 1782
 1783
 1784
 1785
 1786
 1787
 1788
 1789
 1790
 1791
 1792
 1793
 1794
 1795
 1796
 1797
 1798
 1799
 1800
 1801
 1802
 1803
 1804
 1805
 1806
 1807
 1808
 1809
 1810
 1811
 1812
 1813
 1814
 1815
 1816
 1817
 1818
 1819
 1820
 1821
 1822
 1823
 1824
 1825
 1826
 1827
 1828
 1829
 1830
 1831
 1832
 1833
 1834
 1835
 1836
 1837
 1838
 1839
 1840
 1841
 1842
 1843
 1844
 1845
 1846
 1847
 1848
 1849
 1850
 1851
 1852
 1853
 1854
 1855
 1856
 1857
 1858
 1859
 1860
 1861
 1862
 1863
 1864
 1865
 1866
 1867
 1868
 1869
 1870
 1871
 1872
 1873
 1874
 1875
 1876
 1877
 1878
 1879
 1880
 1881
 1882
 1883
 1884
 1885
 1886
 1887
 1888
 1889
 1890
 1891
 1892
 1893
 1894
 1895
 1896
 1897
 1898
 1899
 1900
 1901
 1902
 1903
 1904
 1905
 1906
 1907
 1908
 1909
 1910
 1911
 1912
 1913
 1914
 1915
 1916
 1917
 1918
 1919
 1920
 1921
 1922
 1923
 1924
 1925
 1926
 1927
 1928
 1929
 1930
 1931
 1932
 1933
 1934
 1935
 1936
 1937
 1938
 1939
 1940
 1941
 1942
 1943
 1944
 1945
 1946
 1947
 1948
 1949
 1950
 1951
 1952
 1953
 1954
 1955
 1956
 1957
 1958
 1959
 1960
 1961
 1962
 1963
 1964
 1965
 1966
 1967
 1968
 1969
 1970
 1971
 1972
 1973
 1974
 1975
 1976
 1977
 1978
 1979
 1980
 1981
 1982
 1983
 1984
 1985
 1986
 1987
 1988
 1989
 1990
 1991
 1992
 1993
 1994
 1995
 1996
 1997
 1998
 1999
 2000
 2001
 2002
 2003
 2004
 2005
 2006
 2007
 2008
 2009
 2010
 2011
 2012
 2013
 2014
 2015
 2016
 2017
 2018
 2019
 2020
 2021
 2022
 2023
 2024
 2025
 2026
 2027
 2028
 2029
 2030
 2031
 2032
 2033
 2034
 2035
 2036
 2037
 2038
 2039
 2040
 2041
 2042
 2043
 2044
 2045
 2046
 2047
 2048
 2049
 2050
 2051
 2052
 2053
 2054
 2055
 2056
 2057
 2058
 2059
 2060
 2061
 2062
 2063
 2064
 2065
 2066
 2067
 2068
 2069
 2070
 2071
 2072
 2073
 2074
 2075
 2076
 2077
 2078
 2079
 2080
 2081
 2082
 2083
 2084
 2085
 2086
 2087
 2088
 2089
 2090
 2091
 2092
 2093
 2094
 2095
 2096
 2097
 2098
 2099
 2100
 2101
 2102
 2103
 2104
 2105
 2106
 2107
 2108
 2109
 2110
 2111
 2112
 2113
 2114
 2115
 2116
 2117
 2118
 2119
 2120
 2121
 2122
 2123
 2124
 2125
 2126
 2127
 2128
 2129
 2130
 2131
 2132
 2133
 2134
 2135
 2136
 2137
 2138
 2139
 2140
 2141
 2142
 2143
 2144
 2145
 2146
 2147
 2148
 2149
 2150
 2151
 2152
 2153
 2154
 2155
 2156
 2157
 2158
 2159
 2160
 2161
 2162
 2163
 2164
 2165
 2166
 2167
 2168
 2169
 2170
 2171
 2172
 2173
 2174
 2175
 2176
 2177
 2178
 2179
 2180
 2181
 2182
 2183
 2184
 2185
 2186
 2187
 2188
 2189
 2190
 2191
 2192
 2193
 2194
 2195
 2196
 2197
 2198
 2199
 2200
 2201
 2202
 2203
 2204
 2205
 2206
 2207
 2208
 2209
 2210
 2211
 2212
 2213
 2214
 2215
 2216
 2217
 2218
 2219
 2220
 2221
 2222
 2223
 2224
 2225
 2226
 2227
 2228
 2229
 2230
 2231
 2232
 2233
 2234
 2235
 2236
 2237
 2238
 2239

FIGURE 103

GAGCAGTGTTCCTGCTGGAGCCG**ATG**CCAAAAACCATGCATTTCTTATTCAGATTCATTGTTT
TCTTTTATCTGTGGGGCCTTTTTACTGCTCAGAGACAAAAGAAAGAGGAGAGCACCGAAGAA
GTGAAAATAGAAGTTTTGCATCGTCCAGAAAACCTGCTCTAAGACAAGCAAGAAGGGAGACCT
ACTAAATGCCCATTATGACGGCTACCTGGCTAAAGACGGCTCGAAATTCTACTGCAGCCGGA
CACAAAATGAAGGCCACCCCAAATGGTTTGTTCTTGGTGTTGGGCAAGTCATAAAAGGCCTA
GACATTGCTATGACAGATATGTGCCCTGGAGAAAAGCGAAAAGTAGTTATAACCCCTTCATT
TGCATACGGAAAGGAAGGCTATGCAGAAGGCAAGATTCCACCGGATGCTACATTGATTTTTG
AGATTGAACTTTATGCTGTGACCAAAGGACCACGGAGCATTGAGACATTTAAACAAATAGAC
ATGGACAATGACAGGCAGCTCTCTAAAGCCGAGATAAACCTCTACTTGCAAAGGGAATTTGA
AAAAGATGAGAAGCCACGTGACAAGTCATATCAGGATGCAGTTTTAGAAAGATATTTTTAAGA
AGAATGACCATGATGGTGATGGCTTCATTTCTCCCAAGGAATACAATGTATACCAACACGAT
GAACTA**TAG**CATATTTGTATTTCTACTTTTTTTTTTTTAGCTATTTACTGTACTTTATGTATA
AAACAAAGTCACTTTTCTCCAAGTTGTATTTGCTATTTTTCCCTATGAGAAGATATTTTGA
TCTCCCAATACATTGATTTTGGTATAATAAATGTGAGGCTGTTTTGCAAACCTTAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 104

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65406

<subunit 1 of 1, 222 aa, 1 stop

<MW: 25794, pI: 6.24, NX(S/T): 1

MPKTMHFLFRFIVFFYLWGLFTAQRQKKEESTEEVKIEVLHRPENCSKTSKKGDLLNAHYDG
YLAKDGSKFYCSRTQNEGHPKWFVLGVGQVIKGLDIAMTDMCPGEKRKVVIPPSFAYGKEGY
AEGKIPPDATLIFEIELYAVTKGPRSIETFKQIDMDNDRQLSKAEINLYLQREFEKDEKPRD
KSYQDAVLEDIFKKNDDHGDGFISPKEYNVYQHDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 219-222

N-glycosylation site.

amino acids 45-48

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 87-223, 129-142

EF-hand calcium-binding domain proteins

amino acids 202-214, 195-214

104/249

FIGURE 105

CAGAAATGCAGGGACCATTGCTTCTTCCAGGCCTCTGCTTTCTGCTGAGCCTCTTTGGAGCT
GTGACTCAGAAAACCAAACTTCCTGTGCTAAGTGCCCCCAAATGCTTCCTGTGTCAATAA
CACTCACTGCACCTGCAACCATGGATATACTTCTGGATCTGGGCAGAACTATTACATTCC
CCTTGGAGACATGTAACGCCAGGCATGGTGGCTCGCGCCTGTAATCCCAGTTCTTTGGGAAG
CCAAGGCAGGTGGATCACCTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATAGTGAAAC
CCCGTGTCTACTAAAAATACAAAAATCAGCCGGGCGTGGTGGTGCATGCCTGCAATCCCAGT
TACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCTCAGGAGGCAGAAGTTGCAGTGAACCC
AGATCCTGCCATTGCACTCCAGCATGGATGACAGAGCAAGACTCCGTCTCAAAAAGAAAAGA
TAGTTTCTTGTTTCATTTGCGGACTGCCCTCTCAGTGTTTCCTGGGATCCCCTCCCAAATAA
AGTACTTATATTCTC

105/249

FIGURE 106

MQGPLLLPGLCFLLSLFGAVTQKTKTSCAKCPPNASCVNNTHTCTCNHGYTSGSGQKLFTFPL
ETCNARHGG SRL

Signal peptide:

amino acids 1-18

106/249

FIGURE 107

CAAGCAGGTCATCCCCTTGGTGACCTTCAAAGAGAAGCAGAGAGGGCAGAGGTGGGGGGGCAC
AGGGAAAGGGTGACCTCTGAGATTCCCCTTTTCCCCCAGACTTTGGAAGTGACCCACC**ATGG**
GGCTCAGCATCTTTTGTCTCCTGTGTGTTCTTGGGCTCAGCCAGGCAGCCACACCGAAGATT
TTCAATGGCACTGAGTGTGGGCGTAACTCACAGCCGTGGCAGGTGGGGCTGTTTGAGGGGCAC
CAGCCTGCGCTGCGGGGGTGTCTTATTGACCACAGGTGGGTCTCACAGCGGCTCACTGCA
GCGGCAGCAGGTACTGGGTGCGCCTGGGGGAACACAGCCTCAGCCAGCTCGACTGGACCGAG
CAGATCCGGCACAGCGGCTTCTCTGTGACCCATCCCGGCTACCTGGGAGCCTCGACGAGCCA
CGAGCACGACCTCCGGCTGCTGCGGCTGCGCCTGCCCCGTCCGCGTAACCAGCAGCGTTCAAC
CCCTGCCCCCTGCCCAATGACTGTGCAACCGCTGGCACCGAGTGCCACGTCTCAGGCTGGGGC
ATCACCAACCACCCACGGAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGT
CTCCCATGCCACCTGCCATGGTGTGTATCCCGGGAGAATCACGAGCAACATGGTGTGTGCAG
GCGGCGTCCCGGGGCAGGATGCCTGCCAGGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGGA
GTCCTTCAAGGTCTGGTGTCTGGGGGTCTGTGGGGCCCTGTGGACAAGATGGCATCCCTGG
AGTCTACACCTATATTTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAAC**TGAC**
CTGTTTCCCTCCACCTCCACCCCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACC
AATATCTCCTCCATCACTTCCCCTAGCTCCACTCTTGTTGGCCTGGGAACTTCTTGGAACCTT
TAACTCCTGCCAGCCCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGA
ATAAATATAAATGAAGGAGGGGCAAAAAAAAAAAAAA

FIGURE 108

MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLTAAH
CSGSRYWVRLGEHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLLRLRLPVRVTSSV
QPLPLPNDCATAGTECHVSGWGITNHPRNPFPDLLQCLNLSIVSHATCHGVYPGRITSNMVC
AGGVPGQDACQGDSGGPLVCGGVLQGLVSWGSGVPCGQDGIPGVYTYICKYVDWIRMIMRNN

Signal peptide:

amino acids 1-17

108/249

FIGURE 109

GCGGCCACACGCAGCTAGCCGGAGCCCGGACCAGGCGCCTGTGCCTCCTCCTCGTCCCTCGC
CGCGTCCGCGAAGCCTGGAGCCGGCGGGAGCCCCGCGCTCGCC**ATG**TCGGGCGAGCTCAGCA
ACAGGTTCCAAGGAGGGAAGGCGTTTCGGCTTGCTCAAAGCCCGGCAGGAGAGGAGGCTGGCC
GAGATCAACCGGGAGTTTCTGTGTGACCAGAAGTACAGTGATGAAGAGAACCTTCCAGAAAA
GCTCACAGCCTTCAAAGAGAAGTACATGGAGTTTGACCTGAACAATGAAGGCGAGATTGACC
TGATGTCTTTAAAGAGGATGATGGAGAAGCTTGGTGTCCCCAAGACCCACCTGGAGATGAAG
AAGATGATCTCAGAGGTGACAGGAGGGGTCAGTGACACTATATCCTACCGAGACTTTGTGAA
CATGATGCTGGGGAAACGGTCGGCTGTCCTCAAGTTAGTCATGATGTTTGAAGGAAAAGCCA
ACGAGAGCAGCCCCAAGCCAGTTGGCCCCCCTCCAGAGAGAGACATTGCTAGCCTGCCCC**TGA**
GGACCCCGCCTGGACTCCCCAGCCTTCCACCCCATACCTCCCTCCCGATCTTGCTGCCCTT
CTTGACACACTGTGATCTCTCTCTCTCATTTGTTTGGTCATTGAGGGTTTGTGTTGTGTTT
TCATCAATGTCTTTGTAAAGCACAAATTATCTGCCTTAAAGGGGCTCTGGGTGCGGGGAATCC
TGAGCCTTGGGTCCCCTCCCTCTCTTCTTCCCTCCTTCCCCGCTCCCTGTGCAGAAGGGCTG
ATATCAAACCAAAACTAGAGGGGGCAGGGCCAGGGCAGGGAGGCTTCCAGCCTGTGTTCCC
CTCACTTGGAGGAACCAGCACTCTCCATCCTTTCAGAAAGTCTCCAAGCCAAGTTCAGGCTC
ACTGACCTGGCTCTGACGAGGACCCAGGCCACTCTGAGAAGACCTTGGAGTAGGGACAAGG
CTGCAGGGCCTCTTTCGGGTTTCCTTGGACAGTGCCATGGTTCCAGTGCTCTGGTGTACCCC
AGGACACAGCCACTCGGGGCCCCGCTGCCCCAGCTGATCCCCACTATTCCACACCTCTTCT
CATCCTCAGTGATGTGAAGGTGGGAAGGAAAGGAGCTTGGCATTGGGAGCCCTTCAAGAAGG
TACCAGAAGGAACCCTCCAGTCCTGCTCTCTGGCCACACCTGTGCAGGCAGCTGAGAGGCAG
CGTGCAGCCCTACTGTCCCTTACTGGGGCAGCAGAGGGCTTCGGAGGCAGAAGTGAGGCCTG
GGGTTTGGGGGGAAAGGTCAGCTCAGTGCTGTTCCACCTTTTAGGGAGGATACTGAGGGGGAC
CAGGATGGGAGAATGAGGAGTAAAATGCTCACGGCAAAGTCAGCAGCACTGGTAAGCCAAGA
CTGAGAAATACAAGGTTGCTTGTCTGACCCCAATCTGCTTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 110

MSGELSNRFQGGKAFGLLKAQERRLAELNREFLCDQKYSDEENLPEKLTAFKEKYMEFDLN
NEGEIDLMSLKRMMEKLGVPKTHLEMKKMISEVTGGVSDTISYRDFVNMMLGKRS AVLKLV
MFEGKANESSPKPVGPPPERDIASLP

110/249

FIGURE 111A

CGCGCTCCCCGCGCGCCTCCTCGGGCTCCACGCGTCTTGCCCCGCAGAGGCAGCCTCCTCCA
GGAGCGGGGCCCTGCACACC**ATG**GGCCCCGGGTGGGCAGGGGTTCGGCGCCGCCGTGCGCGCC
CGCTGGCGCTGGCCTTGGCGCTGGCGAGCGTCCTGAGTGGGCCTCCAGCCGTGCGCTGCCC
CACCAAGTGTACCTGCTCCGCTGCCAGCGTGGACTGCCACGGGCTGGGCCTCCGCGCGGTTT
CTCGGGGCATCCCCCGCAACGCTGAGCGCCTTGACCTGGACAGAAATAATATCACCAGGATC
ACCAAGATGGACTTCGCTGGGCTCAAGAACCTCCGAGTCTTGATCTGGAAGACAACCAGGT
CAGCGTCATCGAGAGAGGCGCCTTCCAGGACCTGAAGCAGCTAGAGCGACTGCGCCTGAACA
AGAATAAGCTGCAAGTCCTTCCAGAATTGCTTTTCCAGAGCACGCCGAAGCTCACCAGACTA
GATTTGAGTGAAAACCAGATCCAGGGGATCCCGAGGAAGGCGTTCCGCGGCATCACCGATGT
GAAGAACCTGCAACTGGACAACAACCACATCAGCTGCATTGAAGATGGAGCCTTCCGAGCGC
TGCGCGATTTGGAGATCCTTACCCTCAACAACAACAACATCAGTCGCATCCTGGTCACCAGC
TTCAACCACATGCCGAAGATCCGAACCTCTGCGCCTCCACTCCAACCACCTCTACTGCGACTG
CCACCTGGCCTGGCTCTCGGATTGGCTGCGACAGCGACGGACAGTTGGCCAGTTACACTCT
GCATGGCTCCTGTGCATTTGAGGGGCTTCAACGTGGCGGATGTGCAGAAGAAGGAGTACGTG
TGCCCGAGCCCCCACTCGGAGCCCCCATCCTGCAATGCCAACTCCATCTCCTGCCCTTCGCC
CTGCACGTGCAGCAATAACATCGTGGACTGTGAGGAAAGGGCTTGATGGAGATTCTGCCA
ACTTGCCGGAGGGGCATCGTCGAAATACGCCTAGAACAGAACTCCATCAAAGCCATCCCTGCA
GGAGCCTTCACCCAGTACAAGAACTGAAGCGAATAGACATCAGCAAGAATCAGATATCGGA
TATTGCTCCAGATGCCTTCCAGGGCCTGAAATCACTCACATCGCTGGTCCTGTATGGGAACA
AGATCACCGAGATTGCCAAGGGACTGTTTGTATGGGCTGGTGTCCCTACAGCTGCTCCTCCTC
AATGCCAACAAGATCAACTGCCTGCGGGTGAACACGTTTTCAGGACCTGCAGAACCTCAACTT
GCTCTCCCTGTATGACAACAAGCTGCAGACCATCAGCAAGGGGCTCTTCGCCCCCTCTGCAGT
CCATCCAGACACTCCACTTAGCCCCAAAACCCATTTGTGTGCGACTGCCACTTGAAGTGGCTG
GCCGACTACCTCCAGGACAACCCCATCGAGACAAGCGGGGCCCGCTGCAGCAGCCCCGCGCCG
ACTCGCCAACAAGCGCATCAGCCAGATCAAGAGCAAGAAGTTCCGCTGCTCAGGCTCCGAGG
ATTACCGCAGCAGGTTTCAGCAGCGAGTGCTTCATGGACCTCGTGTGCCCCGAGAAGTGTGCG
TGTGAGGGGCACGATTGTGGACTGCTCCAACCAGAAGCTGGTCCGCATCCCAAGCCACCTCCC
TGAATATGTACCGACCTGCGACTGAATGACAATGAGGTATCTGTTCTGGAGGCCACTGGCA
TCTTCAAGAAGTTGCCAACCTGCGGAAAATAAATCTGAGTAACAATAAGATCAAGGAGGTG
CGAGAGGGAGCTTTCGATGGAGCAGCCAGCGTGCAGGAGCTGATGCTGACAGGGAACCAGCT
GGAGACCGTGCACGGGCGCGTGTTCGCTGGCCTCAGTGGCCTCAAAACCTTGATGCTGAGGA
GTAACCTGATCAGCTGTGTGAGTAATGACACCTTTGCCGGCCTGAGTTCCGGTGAGACTGCTG
TCCCTCTATGACAATCGGATCACCAACCATCACCCCTGGGGCCTTCACCACGCTTGTCTCCCT
GTCCACCATAAACCTCCTGTCCAACCCCTTCAACTGCAACTGCCACCTGGCCTGGCTCGGCA
AGTGGTTGAGGAAGAGGCGGATCGTCAGTGGGAACCCTAGGTGCCAGAAGCCATTTTTCCTC
AAGGAGATTCCCATCCAGGATGTGGCCATCCAGGACTTCACCTGTGATGGCAACGAGGAGAG
TAGCTGCCAGCTGAGCCCGCGCTGCCCGGAGCAGTGCACCTGTATGGAGACAGTGGTGCGAT
GCAGCAACAAGGGGCTCCGCGCCCTCCCCAGAGGCATGCCCAAGGATGTGACCGAGCTGTAC
CTGGAAGGAAACCACCTAACAGCCGTGCCAGAGAGCTGTCCGCCCTCCGACACCTGACGCT
TATTGACCTGAGCAACAACAGCATCAGCATGCTGACCAATTACACCTTCAGTAACATGTCTC
ACCTCTCCACTCTGATCCTGAGCTACAACCGGCTGAGGTGCATCCCCGTCCACGCCTTCAAC
GGGCTGCGGTCCCTGCGAGTGCTAACCCTCCATGGCAATGACATTTCCAGCGTTCTTGAAGG
CTCCTTCAACGACCTCACATCTCTTTCCCATCTGGCGCTGGGAACCAACCCACTCCACTGTG
ACTGCAGTCTTCGGTGGCTGTGCGAGTGGGTGAAGGCGGGGTACAAGGAGCCTGGCATCGCC
CGCTGCAGTAGCCCTGAGCCCATGGCTGACAGGCTCCTGCTCACCACCCCAACCCACCGCTT
CCAGTGCAAAGGGCCAGTGGACATCAACATTTGTGGCCAAATGCAATGCCTGCCTCTCCAGCC
CGTGCAAGAATAACGGGACATGCACCCAGGACCCTGTGGAGCTGTACCGCTGTGCCTGCCCC

TACAGCTACAAGGGCAAGGACTGCACTGTGCCCATCAACACCTGCATCCAGAACCCCTGTCA
GCATGGAGGCACCTGCCACCTGAGTGACAGCCACAAGGATGGGTTCAGCTGCTCCTGCCCTC
TGGGCTTTGAGGGGCAGCGGTGTGAGATCAACCCAGATGACTGTGAGGACAACGACTGCGAA
AACAATGCCACCTGCGTGGACGGGATCAACAACCTACGTGTGTATCTGTCCGCCTAACTACAC
AGGTGAGCTATGCGACGAGGTGATTGACCACTGTGTGCCTGAGCTGAACCTCTGTGAGCATG
AGGCCAAGTGCATCCCCCTGGACAAAGGATTGAGTGCAGTGTGTCCCTGGCTACAGCGGG
AAGCTCTGTGAGACAGACAATGATGACTGTGTGGCCACAAAGTGCCGCCACGGGGCCAGTG
CGTGGACACAATCAATGGCTACACATGCACCTGCCCCAGGGCTTCAGTGGACCCCTTCTGTG
AACACCCCCCACCATGGTCCCTACTGCAGACCAGCCCATGCGACCAGTACGAGTGCCAGAAC
GGGGCCAGTGCATCGTGGTGCAGCAGGAGCCACCTGCCGCTGCCACCAGGCTTCGCCGG
CCCCAGATGCGAGAAGCTCATCACTGTCAACTTCGTGGCAAAGACTCCTACGTGGAAGTGG
CCTCCGCCAAGGTCCGACCCCAGGCCAACATCTCCCTGCAGGTGGCCACTGACAAGGACAAC
GGCATCCTTCTCTACAAAGGAGACAATGACCCCCCTGGCACTGGAGCTGTACCAGGGCCACGT
GCGGCTGGTCTATGACAGCCTGAGTTCCCCTCCAACCACAGTGTACAGTGTGGAGACAGTGA
ATGATGGGCAGTTTCACAGTGTGGAGCTGGTGACGCTAAACCAGACCCTGAACCTAGTAGTG
GACAAAGGAACTCCAAAGAGCCTGGGGAAGCTCCAGAAGCAGCCAGCAGTGGGCATCAACAG
CCCCCTCTACCTTGGAGGCATCCCCACCTCCACCGGCCCTCTCCGCCTTGCGCCAGGGCACGG
ACCGGCCTCTAGGCGGCTTCCACGGATGCATCCATGAGGTGCGCATCAACAACGAGCTGCAG
GACTTCAAGGCCCTCCACCCACAGTCCCTGGGGGTGTACCAAGGCTGCAAGTCTTGCACCGT
GTGCAAGCACGGCCTGTGCCGCTCCGTGGAGAAGGACAGCGTGGTGTGCGAGTGCCGCCAG
GCTGGACCGGCCCACTCTGCGACCAGGAGGCCCGGGACCCCTGCCTCGGCCACAGATGCCAC
CATGGAAAATGTGTGGCAACTGGGACCTCATACATGTGCAAGTGTGCCGAGGGCTATGGAGG
GGACTTGTGTGACAACAAGAATGACTCTGCCAATGCCTGCTCAGCCTTCAAGTGTACCATG
GGCAGTGCCACATCTCAGACCAAGGGGAGCCCTACTGCCTGTGCCAGCCCGGCTTTAGCGGC
GAGCACTGCCAACAAGAGAATCCGTGCCTGGGACAAGTAGTCCGAGAGGTGATCCGCCGCCA
GAAAGGTTATGCATCATGTGCCACAGCCTCCAAGGTGCCCATCATGGAATGTGCTGGGGGCT
GTGGGGCCCCAGTGCTGCCAGCCCACCCGCAGCAAGCGGCGGAAATACGTCTTCCAGTGCACG
GACGGCTCCTCGTTTGTAGAAGAGGTGGAGAGACACTTAGAGTGCGGCTGCCTCGCGTGTTC
CTAAGCCCCTGCCCGCCTGCCTGCCACCTCTCGGACTCCAGCTTGATGGAGTTGGGACAGCC
ATGTGGGACCCCCCTGGTGATTGAGCATGAAGGAAATGAAGCTGGAGAGGAAGGTAAAGAAGA
AGAGAATATTAAGTATATTGTAAATAAACAAAAATAGAACTTAAAAAAAAAAAAAAAAAAAA
AAAAAA

FIGURE 112

MAPGWAGVGAAVRARLALALASVLSGPPAVACPTKCTCSAASVDCHGLGLRAVPRGIPRN
 AERLDLDRNNITRITKMDFAGLKNLRVLHLEDNQVSVIERGAFQDLKQLERLRLNKNKLQVL
 PELLFQSTPKLTRLDLSENQIQGIPRKAFRGITDVKNLQLDNNHISCIEDGAFRALRDLEIL
 TLNNNNISRILVTSFNHMPKIRTLRLHSNHLYCDCHLAWLSDWLRQRRTVGQFTLCMAPVHL
 RGFNVADVQKKEYVCPAPHSEPPSCNANSISCPSPCTCSNNIVDCRGKGLMEIPANLPEGIV
 EIRLEQNSIKAIPAGAFQYKKLKRIDISKQISDIAPDAFQGLKSLTSLVLYGNKITEIAK
 GLFDGLVSLQLLLLNANKINCLRVNTFQDLQNLNLLSLYDNKLQTISKGLFAPLQSIQTLHL
 AQNPFCVCDCHLKWLDYLQDNPIETSGARCSSPRRLANKRISQIKSKKFRCSGSEDIRSRFS
 SECMDLVCPEKCRCEGTIVDCSNQKLVRIPSHLPEYVTDLRLNDNEVSVLEATGIFKKLPN
 LRKINLSNNKIKEVREGAFDGAASVQELMLTGNQLETVHGRVFRGLSGLKTLMLRSNLISCV
 SNTDFAGLSSVRLLSLYDNRIITTITPGAFTTLVSLSTINLLSNPFNCNCHLAWLGKWLKRKR
 IVSGNPRCQKPFFLKEIPIQDVAIQDFTCDGNEESSCQLSPRCPEQCTCMETVVRC SNKGLR
 ALPRGMPKDVTELYLEGNHLTAVPRELSALRHLLTIDLSNNSISMLTNYTFSNMSHLSTLIL
 SYNRLRCIPVHAFNGLRSLRVLTLLHGNDISSVPEGSFNDLTSLSHLALGTNPLHDCSLRWL
 SEWVKAGYKEPGIARCSSPEPMADRLLLTTPTHRFQCKGPVDINIVAKCNACLSPPCKNNGT
 CTQDPVELYRCACPYSYKGKDCTVPINTCIQNPCQHGGTCHLSDSHKDGFSCSCPLGFEGQR
 CEINPDDCEDNDCENNATCVDGINNYVCICPPNYTGELCDEVIDHCVPELNLCQHEAKCIPL
 DKGFSCECVPGYSGKLCETDNDDCVAHKCRHGAQCVDTINGYTCTCPQGFSGPFCEHPPPMV
 LLQTSPCDQYECQNGAQCIVVQQEPTCRCPPGFAGPRCEKLITVNFVGKDSYVELASAKVRP
 QANISLQVATDKDNGILLYKGDNDPLALELYQGHVRLVYDSLSSPPTTVYSVETVNDGQFHS
 VELVTLNQTNLNVVDKGTPKSLGKLQKQPAVGINSPLYLGGIPTSTGLSALRQGTDRPLGGF
 HGCIEHVRINNELQDFKALPPQSLGVSPGCKSCTVCKHGLCRSVEKDSVVCECRPGWTGPLC
 DQEARDPCLGHRCHHGKCVATGTSYMCKCAEGYGGDLCDNKNDSANACSAFKCHHGQCHISD
 QGEPYCLCQPGFSGEHCQQENPCLGQVVREVIRRQKGYASCATASKVPIMECRGGCGPQCCQ
 PTRSKRRKYVFQCTDGSSFVEEVERHLECGCLACS

Signal peptide:

amino acids 1-27

FIGURE 113

GGATGCAGGACGCTCCCCTGAGCTGCCTGTCACCGACTAGGTGGAGCAGTGTTTCTTCCGCA
GACTCAACTGAGAAGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTCAGTTCTGTCT
CCGGCAGGCTTTGAGGATGAAGGCTGCGGGCATTCTGACCCTCATTGGCTGCCTGGTCACAG
GCGCCGAGTCCAAAATCTACACTCGTTGCAAACCTGGCAAATAATTCTCGAGGGCTGGCCTG
GACAATTACTGGGGCTTCAGCCTTGGAACCTGGATCTGCATGGCATATTATGAGAGCGGCTA
CAACACCACAGCCCCGACGGTCCTGGATGACGGCAGCATCGACTATGGCATCTTCCAGATCA
ACAGCTTCGCGTGGTGCAGACGCGGAAAGCTGAAGGAGAACAACCACTGCCATGTCGCCTGC
TCAGCCTTGATCACTGATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTAAAGA
GACACAAGGAATGAACTATTGGCAAGGCTGGAAGAAACATTGTGAGGGCAGAGACCTGTCCG
AGTGGA AAAAAGGCTGTGAGGTTTCCTAAACTGGAACCTGGACCCAGGATGCTTTGCAGCAAC
GCCCTAGGATTTGCAGTGAATGTCCAAATGCCTGTGTCATCTTGTCCCGTTTCCTCCCAATA
TTCCTTCTCAAACCTGGAGAGGGAAAATTAAGCTATACTTTTAAGAAAATAAATATTTCCAT
TTAAATGTC

115/249

FIGURE 114

MKAAGILTLIGCLVTGAESKIYTRCKLAKIFSRAGLDNYWGFSLGNWICMAYYESGYNTTAP
TVLDDGSIDYGIFQINSFAWCRRGKCLKENNHCHVACSALITDDLTDALICARKIVKETQGMN
YWQGWKKHCEGRDLSEWKKGCEVS

Signal peptide:

amino acids 1-19

115/249

FIGURE 115

CAGGCCATTTGCATCCCACTGTCCTTGTGTTTCGGAGCCAGGCCACACCGTCCTCAGCAGTGT
 CATGTGTTAAAAACGCCAAGCTGAATATATC**ATG**CCCCTATTAAACTTGTACATGGCTCCC
 CATTGGTTTTTTGGAGAAAAGTTCAAGCTTTTTACCTTGGTGTCTGCCTGTATCCCAGTGTTT
 AGGCTGGCTAGACGGCGGAAGAAGATCCTATTTTACTGTCACTTCCCAGATCTGCTTCTCAC
 CAAGAGAGATTCTTTTCTTAAACGACTATACAGGGCCCCAATTGACTGGATAGAGGAATACA
 CCACAGGCATGGCAGACTGCATCTTAGTCAACAGCCAGTTCACAGCTGCTGTTTTTAAGGAA
 ACATTCAAGTCCCTGTCTCACATAGACCCTGATGTCCTCTATCCATCTCTAAATGTCACCAG
 CTTTGACTCAGTTGTTTCTGAAAAGCTGGATGACCTAGTCCCCAAGGGGAAAAAATTCCTGC
 TGCTCTCCATCAACAGATACGAAAGGAAGAAAAATCTGACTTTGGCACTGGAAGCCCTAGTA
 CAGCTGCGTGGAAGATTGACATCCCAAGATTGGGAGAGGGTTCATCTGATCGTGGCAGGTGG
 TTATGACGAGAGAGTCCCTGGAGAATGTGGAACATTATCAGGAATTGAAGAAAATGGTCCAAC
 AGTCCGACCTTGGCCAGTATGTGACCTTCTTGAGGTCTTTCTCAGACAAACAGAAAATCTCC
 CTCTCCACAGCTGCACGTGTGTGCTTTACACACCAAGCAATGAGCACTTTGGCATTGTCCC
 TCTGGAAGCCATGTACATGCAGTGCCCAAGTCAATTGCTGTTAATTCGGGTGGACCCTTGGAGT
 CCATTGACCACAGTGTACAGGGTTTCTGTGTGAGCCTGACCCGGTGCACTTCTCAGAAGCA
 ATAGAAAAGTTCATCCGTGAACCTTCCTTAAAAGCCACCATGGGCTGGCTGGAAGAGCCAG
 AGTGAAGGAAAAATTTTCCCCTGAAGCATTTACAGAACAGCTCTACCGATATGTTACCAAAC
 TGCTGGTA**TAA**TCAGATTGTTTTTAAGATCTCCATTAATGTCATTTTTATGGATTGTAGACC
 CAGTTTTGAAACCAAAAAAGAAACCTAGAATCTAATGCAGAAGAGATCTTTTAAAAAATAAA
 CTTGAGTCTTGAATGTGAGCCACTTTCCTATATACCACACCTCCCTGTCCACTTTTCAGAAA
 AACCATGTCTTTTATGCTATAATCATTCCAAATTTTGCCAGTGTTAAGTTACAAATGTGGTG
 TCATTCCATGTTTCAGCAGAGTATTTTAATTATATTTTCTCGGGATTATTGCTCTTCTGTCTA
 TAAATTTTGAATGATACTGTGCCTTAATTGGTTTTTCATAGTTTAAGTGTGTATCATTATCAA
 AGTTGATTAATTTGGCTTCATAGTATAATGAGAGCAGGGCTATTGTAGTTCCCAGATTCAAT
 CCACCGAAGTGTTCACTGTCATCTGTTAGGGAATTTTTGTTTGTCTGTCTTTGCCTGGATC
 CATAGCGAGAGTGCTCTGTATTTTTTTTAAAGATAATTTGTATTTTTTGCACACTGAGATATAA
 TAAAAGGTGTTTATCATAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MPLLKLVHGSPLVFGEKFKLFTLVSAIPVFRLARRRKKILFYCHFDPDLLLTKRDSFLKRLY
RAPIDWIEEYTTGMADCILVNSQFTAAVFKETFKSLSHIDPDVLYPSLNVTSFDSVVPEKLD
DLVPKGKKFLLLSINRYERKKNLTLALEALVQLRGRLTSQDWERVHLIVAGGYDERVLENVE
HYQELKKMVQQSDLGQYVTFILRSFSDKQKISLLHSCTCVLYTPSNEHFGIVPLEAMYMQCPV
IAVNSGGPLESIDHSVTGFLCEPDPVHFSEAIEKFIREPSLKATMGLAGRARVKEKFSPEAF
TEQLYRYVTKLLV

Signal peptide:

amino acids 1-15

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 117

GACTACGCCGATCCGAGACGTGGCTCCCTGGGCGGCAGAACC**ATG**TTGGACTTCGCGATCTT
 CGCCGTTACCTTCTTGCTGGCGTTGGTGGGAGCCGTGCTCTACCTCTATCCGGCTTCCAGAC
 AAGCTGCAGGAATTCCAGGGATTACTCCAACCTGAAGAAAAAGATGGTAATCTTCCAGATATT
 GTGAATAGTGGAAGTTTGCATGAGTTCCTGGTTAATTTGCATGAGAGATATGGGCCTGTGGT
 CTCCTTCTGGTTTGGCAGGCGCCTCGTGGTTAGTTTGGGCACTGTTGATGTACTGAAGCAGC
 ATATCAATCCCAATAAGACATCGGACCCTTTTGAAACCATGCTGAAGTCATTATTAAGGTAT
 CAATCTGGTGGTGGCAGTGTGAGTGAAAACACATGAGGAAAAAATTGTATGAAAATGGTGT
 GACTGATTCTCTGAAGAGTAACTTTGCCCTCCTCCTAAAGCTTTCAGAAGAATTATTAGATA
 AATGGCTCTCCTACCCAGAGACCCAGCACGTGCCCCCTCAGCCAGCATATGCTTGGTTTTGCT
 ATGAAGTCTGTTACACAGATGGTAATGGGTAGTACATTTGAAGATGATCAGGAAGTCATTCTG
 CTTCCAGAAGAATCATGGCACAGTTTGGTCTGAGATTGGAAGGCTTTCTAGATGGGTCTAC
 TTGATAAAAACATGACTCGGAAAAACAATATGAAGATGCCCTCATGCAACTGGAGTCTGTT
 TTAAGGAACATCATAAAGAACGAAAAGGAAGGAACCTCAGTCAACATATTTTTCATTGACTC
 CTTAGTACAAGGGAACCTTAATGACCAACAGATCCTAGAAGACAGTATGATATTTTCTCTGG
 CCAGTTGCATAATAACTGCAAAATTGTGTACCTGGGCAATCTGTTTTTTAACCACCTCTGAA
 GAAGTTCAAAAAAATTATATGAAGAGATAAACCAAGTTTTTTGAAATGGTCCTGTTACTCC
 AGAGAAAATTGAGCAGCTCAGATATTGTCAGCATGTGCTTTGTGAAACTGTTGCAACTGCCA
 AACTGACTCCAGTTTCTGCCCAGCTTCAAGATATTGAAGGAAAAATTGACCGATTTATTATT
 CCTAGAGAGACCCTCGTCCTTTATGCCCTTGGTGTGGTACTTCAGGATCCTAATACTTGGCC
 ATCTCCACACAAGTTTGATCCAGATCGGTTTGATGATGAATTAGTAATGAAAACTTTTTCCT
 CACTTGGATTCTCAGGCACACAGGAGTGTCCAGAGTTGAGGTTTGCATATATGGTGACCACA
 GTACTTCTTAGTGTATTGGTGAAGAGACTGCACCTACTTTCTGTGGAGGGACAGGTTATTGA
 AACAAAGTATGAACTGGTAACATCATCAAGGGAAGAAGCTTGGATCACTGTCTCAAAGAGAT
 AT**TAA**ATTTTATACATTTAAATCATTTGTTAAATTGATTGAGGAAAACAACCATTTAAAAA
 AAATCTATGTTGAATCCTTTTATAAACAGTATCACTTTGTAATATAAACACCTATTTGTAC
 TTAA

FIGURE 118

MLDFAIFAVTFLLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLPDIVNSGSLHEFLVNLH
ERYGPVVSFWFGRRLLVSLGTVDLKQHINPNKTSDFETMLKSLRYQSGGGSVSENHMRK
KLYENGVTDSLKSNFALLLLKLSEELLDKWLSYPETQHVPLSQHMLGFAMKSVTQMVMGSTFE
DDQEVIRFQKNHGTWVSEIGKGFLDGS�DKNMTRKKQYEDALMQLESVLRNIIKERKGRNFS
QHIFIDSLVQGNLNDQQILEDSMIFSLASCIITAKLCTWAICFLTTSSEEVQKKLYEEINQVF
GNGPVTPEKIEQLRYCQHVLCEVVRTAKLTPVSAQLQDIEGKIDRFIIPRETLLVLYALGVVL
QDPNTWPSPHKFDPDRFDDELVMKTFSSLGFSGTQECPELRFAYMVTTVLLSVLVKRLHLLS
VEGQVIETKYELVTSSREEAWITVSKRY

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 271-290

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 119

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGA
GACCGCCGCCCTTGTCCTCCGAGGGGCCATGGGCCGGGTCTCAGGGCTTGCGCCCTCTCGCTTC
CTGACGCTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACAT
ACAGGCCTGCCCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTTCAGCTGG
TGGCCGCGCTCTCTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGA
GTCTCCATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGT
GGCCCTGTCCTTCTTCATATTCGAGCGTTGGGAGTGCCTACGTATTGGTACATTTTTGTCT
TCTGCAGTGCCCTTCCAGCTGTCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAA
AAGAAACCCTTCTGATTACCTTCATGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGG
CCGCTTCGTATTCCTGGAAGAAGGAAGGCATAGGCTTCGGTTTTTCCCCTCGGAAACTGCTTC
TGCTGGAGGATATGTGTTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATTGTATTTA
GTGCTTTGTAATAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACA
ATTAAAAAAAAAAAAA

FIGURE 121

TCCCGGACCCTGCCGCCCTGCCACTATGTCCCGCCGCTCTATGCTGCTTGCCTGGGCTCTCC
CCAGCCTCCTTCGACTCGGAGCGGCTCAGGAGACAGAAGACCCGGCCTGCTGCAGCCCCATA
GTGCCCCGGAACGAGTGGAAGGCCCTGGCATCAGAGTGCGCCCAGCACCTGAGCCTGCCCTT
ACGCTATGTGGTGGTATCGCACACGGCGGGCAGCAGCTGCAACACCCCCGCCTCGTGCCAGC
AGCAGGCCCCGGAATGTGCAGCACTACCACATGAAGACACTGGGCTGGTGCGACGTGGGCTAC
AACTTCCTGATTGGAGAAGACGGGCTCGTATACGAGGGCCGTGGCTGGAACTTCACGGGTGC
CCACTCAGGTCACCTTATGGAACCCCATGTCCATTGGCATCAGCTTCATGGGCAACTACATGG
ATCGGGTGCCACACCCCAGGCCATCCGGGCAGCCCAGGGTCTACTGGCCTGCGGTGTGGCT
CAGGGAGCCCTGAGGTCCAACCTATGTGCTCAAAGGACACCGGGATGTGCAGCGTACACTCTC
TCCAGGCAACCAGCTCTACCACCTCATCCAGAATTGGCCACACTACCGCTCCCCCTGAGGGCC
CTGCTGATCCGCACCCCATTCTCCTCCCCTCCCATGGCCAAAACCCCACTGTCTCCTTCTCCA
ATAAAGATGTAGCTC

CCAGCCTCCTTCGACTCGGAGCGGCTCAGGAGACAGAAGACCCGGCCTGCTGCAGCCCCATA

FIGURE 122

MSRRSMLLAWALPSLLRLGAAQETEDPACCSPIVPRNEWKALASECAQHLSLPLRYVVVSHT
AGSSCNTPASCQQQARNVQHYHMKTLGWCDVGYNFLIGEDGLVYEGRGWNFTGAHSGHLWNP
MSIGISFMGNYMDRVPTPQAIRAAQGLLACGVAQGALRSNYVLKGHRDVQRTLSPGNQLYHL
IQNWPHYRSP

Signal peptide:

amino acids 1-20

123/249

FIGURE 123

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAG**ATG**CAACT
 GACTCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGG
 ATGATGGTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCC
 CGGGTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCT
 CCTAGGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCC
 CGAACCACAGCCCCCACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTAC
 TCCAACATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAA
 TGGGACCTTCAGCGTCCACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCC
 TCGTGCCCCCAGTAAAGCTGTAGAGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAG
 GCCTCCAAAATCTTCAACTGCCGGATGGAGTGGGAGAAGGTAGAACGGGGCCGCCGGACCTC
 GCTTTGCACCCACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCAGCCACCT
 GGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTAT
 CGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTACCATAGTGATACCCCTACTACCCATC
 TGGG**TGA**ACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGC
 AGGAGACCATCTGGACACCCGGGCAGGGAAGGGGTGGGCCTCAGGCAGGGAGGGGGGTGGAG
 ACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAGAGAAAGGGTCCCAAGTG
 CTGGTCCCAACCTGAAGCTGTGGAGTGAAGTAGATCACAGGAGCACTGGAGGAGGAGTGGGCT
 CTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCC
 TGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGTTT
 CTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACAGCCTG
 TCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGGC
 CAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATG
 GCATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTC
 CAGCCAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATT
 GCTGATGGCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCC
 ACAGCCCATCCGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAG
 CATCCATGTCCCGGAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCC
 CGGATCTGGATGGCGCCGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGA
 GCATGTGCTGGATCTGTTCTGTGTGTCTGTCTGTGGGTGGGGGGAGGGGAGGGAAGTCTTGT
 GAAACCGCTGATTGCTGACTTTTGTGTGAAGAATCGTGTTCTTGAGCAGGAAATAAGCTT
 GCCCCGGGGCA

FIGURE 124

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66521

><subunit 1 of 1, 252 aa, 1 stop

><MW: 28127, pI: 8.91, NX(S/T): 5

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMAN
STLLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVD
HGNGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGR
RTSLCTHDPAKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVCPCDYNHSDTPY
YPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 125

GTGAATGTGAGGGTTTGATGACTTTCAGATGTCTAGGAACCAGAGTGGGTGCAGGGGCCCCA
 GGCAGGGCTGATTCTTGGGCGGAGGAGAGTAGGGTAAAGGGTTCTGCATGAGCTCCTTAAAG
 GACAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGACTTTGACTTCAAGCCACAGAAT
 TGGTGGAAGTGTGCGCGCCGCCGCCGCTCGCTCCTGCAGCGCTGTCGACCTAGCCGCTAG
 CATCTTCCCGAGCACCGGGATCCCGGGGTAGGAGGCGACGCGGGCGAGCACCAGCGCCAGCC
 GGCTGCGGCTGCCCACACGGCTCACC**ATG**GGCTCCGGGCGCCGGGCGCTGTCCGCGGTGCCG
 GCCGTGCTGCTGGTCCTCACGCTGCCGGGGCTGCCCCTCTGGGCACAGAACGACACGGAGCC
 CATCGTGCTGGAGGGCAAGTGTCTGGTGGTGTGCGACTCGAACCCGGCCACGGACTCCAAGG
 GCTCCTCTTCCCTCCCCGCTGGGGATATCGGTCCGGGCGGCCAACTCCAAGGTCGCCTTCTCG
 GCGGTGCGGAGCACCAACCACGAGCCATCCGAGATGAGCAACAAGACGCGCATCATTTACTT
 CGATCAGATCCTGGTGAATGTGGGTAAATTTTTTTCACATTGGAGTCTGTCTTTGTAGCACCAA
 GAAAAGGAATTTACAGTTTTCAGTTTTCACGTGATTAAAGTCTACCAGAGCCAACTATCCAG
 GTTAACTTGATGTTAAATGGAAGAACAGTAATATCTGCCTTTGCGGGGGACAAAGATGTTAC
 TCGTGAAGCTGCCACGAATGGTGTCTGCTCTACCTAGATAAAGAGGATAAGGTTTACCTAA
 AACTGGAGAAAGGTAATTTGGTTGGAGGCTGGCAGTATTCCACGTTTTCTGGCTTTCTGGTG
 TTCCCCCTA**TAG**GATTCAATTTCTCCATGATGTTTCATCCAGGTGAGGGATGACCCACTCCTG
 AGTTATTGGAAGATCATTTTTTTCATCATTGGATTGATGTCTTTTATTGGTTTCTCATGGGTG
 GATATGGATTCTAAGGATTCTAGCCTGTCTGAACCAATACAAAATTTACAGATTATTTGTG
 TGTGTCTGTTTTCAGTATATTTGGATTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTA
 ACAGTCAAAAGCTGTCTGCAAGACTTATTCTGAATTTCAATTCCTGGGATTACTGAATTAGT
 TACAGATGTGGAATTTTATTTGTTTAGTTTTTAAAGACTGGCAACCAGGTCTAAGGATTAGA
 AAATCTAAAGTTCTGACTTCAATCAACGGTTAGTGTGATACTGCCAAAGAACTGTATACTG
 TGTTAATATATTGATTATATTTGTTTTTATTCCTTTGGAATTAGTTTGTTGGTTCTTGTA
 AAACTTGGAATTTTTTTTTTTCAGTAACTGGTATTATGTTTTCTCTTAAATAAGGTAATGAA
 TGGCTTGCCACAAATTTACCTTGACTACGATATCATCGACATGACTTCTCTCAAAAAAAAAA
 GAATGCTTCATAGTTGTATTTTAATTGTATATGTGAAAGAGTCATATTTTCCAAGTTATATT
 TTCTAAGAAGAAGAATAGATCATAAATCTGACAAGGAAAAAGTTGCTTACCCAAAATCTAAG
 TGCTCAATCCCTGAGCCTCAGCAAAACAGCTCCCCTCCGAGGGAAATCTTATACTTTATTGC
 TCAACTTTAATTAAATGATTGATAATAACCACTTTATTAAAAACCTAAGGTTTTTTTTTTT
 TCCGTAGACATGACCACTTTATTAACTGGTGGTGGGATGCTGTTGTTTCTAATTATACCTAT
 TTTTCAAGGCTTCTGTTGTATTTGAAGTATCATCTGGTTTTGCCTTAACTCTTTAAATTGTA
 TATATTTATCTGTTTAGCTAATATTAAATTCAAATATCCCATATCTAAATTTAGTGCAATAT
 CTTGTCTTTTGTATAGGTCATATGAATTCATAAAATTATTTATGTCTGTTATAGAATAAAGA
 TTAATATATGTTAAAAAA

MGSGRRALSAVPAVLLVLTLPGLPVWAQNDTEPIVLEGKCLVVCDSNPATDSKGSSSSPLGI
SVRAANSKVAFSAVRSTNHEPSEMSNKTRIIYFDQILVNVGNFFTLSEVFVAPRKGIIYSFSF
HVIKVYQSQTIQVNLMLNGKPVISAFAGDKDVTREAAATNGVLLYLDKEDKVYLLKLEKGNLVG
GWQYSTFSGFLVFPL

amino acids 1-27

FIGURE 127

CGGTGGCC**ATG**ACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTC
GCCCTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGC
TTTCTTCTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTA
TTGACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTC
TATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTT
GAAGAGTATAAACCCAGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCT
TGGGCTTTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGG
CCAGGCACAGTGGGCATTCATGGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGAC
GCTGGTCATTATCTTGCTGCATGTATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGA
AAAAGTGGGGCATCCTCCTTATCGTTCTCCTGACCCACCTGCTGGTGTGAGCCAGACCTTC
ATAAGTTCTTATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGCTCATGGGCAC
CTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTCTGCCAAG
ACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGAT**TAA**CCTCAGGGAACCAGCACTTCCCAA
ACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCTTTTTCTGAAAATCCCTTTTTCTG
GTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 128

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66658

><subunit 1 of 1, 257 aa, 1 stop

><MW: 28472, pI: 9.33, NX(S/T): 0

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDN
KDGPTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGF
GIMSGVFSFVNTLSDSLGPSTVGIHGDSPOFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKKW
GILLIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLLCQDKN
FLLYNQSR

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 129

CGGCAACCAGCCGCCGCCACCACCGCTGCCACTGCCGCCCTGCCGGGGCCATGTTTCGCTCTGGGCTTGCCCTTCT
 TGGTGCTCTTGGTGGCCTCGGTGAGAGCCATCTGGGGGTTCTGGGGCCCAAGAACGTCTCGCAGAAAGACGCCG
 AGTTTGAGCGCACCTACGTGGACGAGGTCAACAGCGAGCTGGTCAACATCTACACCTTCAACCATACTGTGACCC
 GCAACAGGACAGAGGGCGTGCCTGTGTCTGTGAACGTCTGAACAAGCAGAAGGGGGCGCCGTTGCTGTTTGTGG
 TCCGCCAGAAGGAGGCTGTGGTGTCTTCCAGGTGCCCTAATCCTGCGAGGGATGTTTCAGCGCAAGTACCTCT
 ACCAAAAAGTGAACGAACCCTGTGTGAGCCCCCACCAGAATGAGTCGGAGATTCAGTTCTTCTACGTGGATG
 TGTCCACCCTGTCAACAGTCAACACCACATACCAGCTCCGGGTGAGCCGATGGACGATTTTGTGCTCAGGACTG
 GGGAGCAGTTGAGCTTCAATACCACAGCAGCAGCCCCAGTACTTCAAGTATGAGTTCCCTGAAGGCGTGGACT
 CGGTAATTGTCAAGGTGACCTCCAACAAGGCCCTTCCCTGCTCAGTCATCTCCATTCAGGATGTGCTGTGTCTG
 TCTATGACCTGGACAACAACGTAGCCTTCATCGGCATGTACCAGACGATGACCAAGAAGGCGGCCATCACCGTAC
 AGCGCAAAGACTTCCCCAGCAACAGCTTTTATGTGGTGGTGGTGAAGACCGAAGACCAAGCCTGCGGGGGCT
 CCCTGCCTTTCTACCCCTTCGCAGAAGATGAACCGGTGATCAAGGGCACC GCCAGAAAACCTGTGAGTGCTGG
 TGTCTCAAGCAGTCAGCTGAGGCATACGTGAGTGGGATGCTCTTTTGCTGGGTATATTTCTCTCTTTTACC
 TGCTGACCGTCTCTGCTGGCCTGCTGGGAGAAGTGGAGGCAAGAAGAAGACCTGCTGGTGGCCATTGACCGAG
 CCTGCCCAGAAAGCGGTACCCCTCGAGTCTGGCTGATTCTTTTCTGGCAGTTCCCTTATGAGGGTTACAAC
 ATGGCTCCTTTGAGAATGTTTCTGGATCTACCGATGGTCTGGTTGACAGCGCTGGCACTGGGGACCTCTCTTACG
 GTTACCAGGGCCGCTCCTTTGAACCTGTAGGTACTCGGCCCGAGTGGACTCCATGAGCTCTGTGGAGGAGGATG
 ACTACGACACATTGACCGACATCGATTCCGACAAGAATGTGATTTCGCACCAAGCAATACCTCTATGTGGCTGACC
 TGGCACGGAAGGACAAGCGTGTCTGCGGAAAAAGTACCAGATCTACTTCTGGAACATTGCCACCATTGCTGTCT
 TCTATGCCCTTCTGTGGTGCAGCTGGTGATCACTACCAGCGGTGGTGAATGTGACAGGGAATCAGGACATCT
 GCTACTACAACCTTCTCTGCGCCCCACCACTGGGCAATCTCAGCGCCTTCAACAACATCCTCAGCAACCTGGGGT
 ACATCCTGTGCGGGCTGCTTTTCTGCTCATCATCCTGCAACGGGAGATCAACCACAACCGGGCCCTGCTGCGCA
 ATGACCTCTGTGCGCTGGAATGTGGGATCCCCAAACACTTTGGGCTTTTCTACGCCATGGGCACAGCCCTGATGA
 TGGAGGGGCTGCTCAGTGCTTGTATCATGTGTGCCCCAACTATACCAATTTCCAGTTTGACACATCGTTTCATGT
 ACATGATCGCCGACTCTGATGCTGAAGCTCTACCAGAAGCGGCACCCGGACATCAACGCCAGCGCCTACAGTG
 CCTACGCTGCTGCGCATTTGTGCTCTTCTCTGTGCTGGGCGTGGTCTTTGGCAAAGGGAACACGGCGTCTCT
 GGATCGTCTTCTCCATCATTCACATCATCGCCACCCTGCTCCTCAGCAGCAGCTCTATTACATAGGGCCGTGGA
 AACTGGACTCGGGGATCTTCCGCCGATCCTCCACGTGCTCTACACAGACTGCATCCGGCAGTGCAGCGCGCCG
 TCTACGTGGACCGCATGGTGTGCTGGTGCATGGGCAACGTGCATCAACTGGTCTGCTGGCTGCCATGGGCTTATCA
 TGCGCCCCAATGATTTTCGCTTCTACTTGTGGCCATTGGCATCTGCAACCTGCTCCTTTACTTTCGCTTCTACA
 TCATCATGAAGCTCCGAGTGGGAGAGGATCAAGCTCATCCCCCTGCTCTGCATCGTTTGACCTCCGTGGTCT
 GGGGCTTCGCGCTCTTCTTCTTCTTCCAGGACTCAGCACCTGGCAGAAAACCCCTGCAGAGTCGAGGGAGCACA
 ACCGGGACTGCATCCTCCTCGACTTCTTTGACGACCACGACATCTGGCACTTCTCTCCTCCATCGCCATGTTCG
 GGTCTTCTGGTGTGCTGACACTGGATGACGACCTGGATACTGTGCAGCGGGACAAAGATCTATGCTTCTAGC
 AGGAGCTGGGCCCTTCGCTTACCTCAAGGGGCCCTGAGCTCCTTTGTGTGATAGACCGGTCACTCTGTCTGCT
 GTGGGGATGAGTCCCAGCACCGCTGCCAGCACTGGATGGCAGCAGGACAGCCAGGTCTAGCTTAGGCTTGCCCT
 GGGACAGCCATGGGGTGGCATGGAACCTTGCAGCTGCCCTTGCCGAGGAGCAGGCCCTGCTCCCCCTGGAACCCCC
 AGATGTTGGCCAAATTGCTGCTTTCTTCTCAGTGTTGGGGCCTTCCATGGGCCCTGTCTTTGGCTCTCCATTT
 GTCCCTTTGCAAGAGGAAGGATGGAAGGGACACCCTCCCCATTTTCATGCTTGCATTTTGGCCGTCTCTCCCC
 ACAATGCCCCAGCCTGGGACCTAAGGCCTCTTTTCTCTCCATACTCCCACTCCAGGGCCTAGTCTGGGGCCTGA
 ATCTCTGTCTGTATCAGGGCCCCAGTTCTCTTTGGGCTGTCCCTGGCTGCCATCACTGCCCATTCAGTCAAGCC
 AGGATGGATGGGGGTATGAGATTTTGGGGGTGGCCAGCTGGTGCCAGACTTTTGGTGCTAAGGCCTGCAAGGGG
 CCTGGGGCAGTGCGTATTCTCTTCCCTCTGACCTGTGCTCAGGGCTGGCTCTTTAGCAATGCGCTCAGCCCAATT
 TGAGAACCGCCTTCTGATTCAAGAGGCTGAATTCAGAGGTACCTCTTTCATCCCATCAGCTCCAGACTGATGCC
 AGCACCAGGACTGGAGGGAGAAGCGCCTCACCCCTTCCCTTCTTTCAGGCCCCCTAGTCTTGCCAAACCCC
 AGCTGGTGGCCTTTTCAGTGCCATTGACACTGCCAAGAATGTCCAGGGGCAAAGGAGGGATGATACAGAGTTCAG
 CCCGTTCTGCCTCCACAGCTGTGGGCACCCCACTGCTTACCTTAGAAAGGGGCTTCAGGAAGGGATGTGCTGTTT
 CCTCTACGTGCCAGTCTTAGCCTCGCTCTAGGACCCAGGGCTGGCTTCTAAGTTTCCGTCCAGTCTTCAGGCA
 AGTTCTGTGTTAGTCATGCACACACATACCTATGAAACCTTGGAGTTTACAAAGAATTGCCCCAGCTCTGGGCAC
 CCTGGCCACCCTGGTCTTGGATCCCCCTTCGTCCCACCTGGTCCACCCAGATGCTGAGGATGGGGGAGCTCAGG
 CGGGGCTCTGCTTTGGGGATGGGAATGTGTTTTCTCCAACTTGTTTTTATAGCTCTGCTTGAAGGGCTGGG
 AGATGAGGTGGGTCTGGATCTTTTCTCAGAGCGTCTCCATGCTATGGTTGCATTTCCGTTTTCTATGAATGAATT
 TGCATTCAATAAACAACCAGACTCAAAAAAAAAAAAAA

FIGURE 130

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66659
><subunit 1 of 1, 832 aa, 1 stop
><MW: 94454, pI: 6.94, NX(S/T): 12
MFALGLPFLVLLVASVESHLGVLGPKNVSQKDAEFERTYVDEVNSELVNIYTFNHTVTRNRT
EGVRVSVNVLNKQKGAPLLFVVRQKEAVVSFQVPLILRGMFQRKYLYQKVERTLCQPPTKNE
SEIQFFYVDVSTLSPVNTTYQLRVSRMDDFVLRTGEQFSFNTTAAQPQYFKYEFPEGVDSVI
VKVTSNKAFPSCSVISIQDVLCPVYDLNNAVFIGMYQMTKKAAITVQRKDFPSNSFYVVVV
VKTEDQACGGSLPFYPFAEDEPVDQGHRQKTL SVLVSQAVTSEAYVSGMLFCLGIFLSFYLL
TVLLACWENWRQKKKTLVAIDRACPESGHPRLADSFPGSSPYEGYNYGSFENVSGSTDGL
VDSAGTGDLSYGYQGRSFEPVGTRPRVDSMSSVEEDDYDTLTDIDSDKNVIRTKQYLYVADL
ARKDKRVLRRKKYQIYFWNIATIAVFYALPVVQLVITYQTVVNVVTGNQDICYNFLCAHPLGN
LSAFNNILSNLGYILLGLLFLLIILQREINHNRALLRNDLCALECGIPKHFGFLFYAMGTALM
MEGLLSACYHVCNPYTNFQFDTSEFMYMIAGLCMLKLYQKRHPDINASAYSAYACLAIVIFFS
VLGVVFGKGNTAFWIVFSIIHIIATLLLSTQLYMGWRWKLDSGIFRRILHVLYTDCIRQCSG
PLYVDRMVLLVMGNVINWSLAAYGLIMRPNDFASYLLAIGICNLLLYFAFYIIMKLRSGERI
KLIPLLCIVCTSVVWGFALFFFFQGLSTWQKTPAESREHNRDCILLDFDDHDIWHFLSSIA
MFGSFLVLLTLDDDLDTVQRDKIYVF
```

Important features of the protein:**Signal peptide:**

amino acids 1-18

Transmembrane domains:

amino acids 292-317, 451-470, 501-520, 607-627, 751-770

Leucine zipper pattern.

amino acids 497-518

N-glycosylation sites.

amino acids 27-30, 54-57, 60-63, 123-126, 141-144, 165-168, 364-367, 476-479, 496-499, 572-575, 603-606, 699-702

FIGURE 131

GCTCTCAAGTGCCTTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGC
 TCTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTG
 ACC**ATG**GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGC
 CCAGCCTGCAGAGCTGTCTGTGGAAGTTCCAGAAAACCTATGGTGGAATTTCCCTTTATAC
 TGACCAAGTTGCCGCTGCCCCGTGAGGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCA
 GGCAAGGCAACTGAGGGCCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAG
 GGCCCTGGACCGAGAGGAGCAGGCAGAGTACCAGCTACAGGTACCCTGGAGATGCAGGATG
 GACATGTCTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTG
 CCCCATTTCTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCC
 CTTCTCTTCCCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACTCGGATCTTCGAT
 TCCACATCCTGAGCCAGGCTCCAGCCCAGCCTTCCCCAGACATGTTCCAGCTGGAGCCTCGG
 CTGGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGCACCAGCCTTGACCACGCCCTGGAGAGGAC
 CTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCAGGCCACTG
 CCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGTCCCTAGAGCCTATCCACCTGGCA
 GAGAATCTCAAAGTCTTATACCCGCACCACATGGCCCAGGTACACTGGAGTGGGGGTGATGT
 GCACTATCACCTGGAGAGCCATCCCCCGGACCCTTTGAAGTGAATGCAGAGGGAAACCTCT
 ACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAG
 AATTCCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTGATGGATGAGAA
 TGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCTCAGTCCAC
 CAGGTACTGAAGTGACTAGACTGTCAGCAGAGGATGCAGATGCCCCCGGCTCCCCCAATTCC
 CACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTCCA
 GGTGGAACCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACA
 TCCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGT
 GAAGTGAAGTCGAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTCCAGAT
 TGGGCCTATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCA
 TTGATGCTGACCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACA
 GAAGGGACTTTTGGCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAA
 GAACCTCAGTTATGAGGCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGA
 AGCTGGTGGGGCCAGGCCCAGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGA
 GTGATGCCACCCCCCAAGTTGGACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCC
 AGCCGGCTCTTTCCCTGCTGACCATCCAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCT
 CCCTAGTCAATGACTCAGAGGGCTGGCTCTGCATTGAGAAATTCTCCGGGGAGGTGCACACC
 GCCCAGTCCCTGCAGGGCGCCCAGCCTGGGGACACCTACACGGTGCTTGTGGAGGGCCAGGA
 TACAGCCCTGACTCTTGCCCCCTGTGCCCTCCCAATACCTCTGCACACCCCGCCAAGACCATG
 GCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCTGGCCAGTGGGCACGGTCCCTACAGC
 TTCACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGCGCCTCCAGACTCTCAATGGTTC
 CCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGTGAACACATAATCCCCGTGG
 TGGTCAGCCACAATGCCCAGATGTGGCAGCTCCTGGTTTCGAGTGATCGTGTGTGCTGCAAC
 GTGGAGGGGCAGTGCATGCGCAAGGTGGGCCGCATGAAGGGCATGCCACGAAGCTGTGGGC
 AGTGGGCATCCTTGTTAGGCACCCTGGTAGCAATAGGAATCTTCTCATCCTCATTTTACACC
 ACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCCTGAAGGCG
 ACTGTC**TGA**ATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAGTCCC
 CTGGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCTCCAT
 CTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAAC
 TTTATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTTTTGCCCAATAATAAAGCCCCA
 GAGAACTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 132

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGD SG
KATEGPFAMDPDSGFLLVTRALDREEQA EYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVP
HFSQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDPMFQLEPRL
GALALSPKGSTSLDHALERTYQLLVQVKMDGQASGHQATATVEVSIIESTWVSLEPIHLAE
NLKVLYPHMAQVHWSSGGDVHYHLESHPPGPF EVNAEGNLYVTRELDREAQA EYLLQVRAQN
SHGEDYAAPLELHVLVMDENDNVPICPPRDPTVSIPELSPPGTEVTRL SAEDADAPGSPNSH
VVYQLLSPEPEDGVEGRAFQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCE
VEVAVTDINDHAPEFITSQIGPISLPEDVEPGTLVAMLT AIDADLEPAFRLMDFAIERGDTE
GTFGLDWE PD SGHVRLRLCKNLSYEAPSHEVVVVVQSVAKLVGPGPGPGATATVTVLVERV
MPPPKLDQESYEASVPISAPAGSFLLTIQSPDPI SRTLRFSLVNDSEGWLCIEKFSGEVHTA
QSLQGAQPGDTYTVLVEAQDTALT LAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSF
TLGPNPTVQRDWRLQTLNGSHAYLT LALHWVEPREHIIPVVVSHNAQMWQLLV RVIVCRCNV
EGQCMRKVGRMKGMPTKLSAVGILVGT LVAIGIFLILIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 133

CCGGGGAC**ATG**AGGTGGATACTGTTCAATTGGGGCCCTTATTGGGTCCAGCATCTGTGGCCAA
 GAAAAATTTTTTGGGGACCAAGTTTTGAGGATTAATGTCAGAAATGGAGACGAGATCAGCAA
 ATTGAGTCAACTAGTGAATTCAAACAACCTGAAGCTCAATTTCTGGAAATCTCCCTCCTCCT
 TCAATCGGCCTGTGGATGTCCTGGTCCCATCTGTCAGTCTGCAGGCATTTAAATCCTTCCTG
 AGATCCCAGGGCTTAGAGTACGCAGTGACAATTGAGGACCTGCAGGCCCTTTTAGACAATGA
 AGATGATGAAATGCAACACAATGAAGGGCAAGAACGGAGCAGTAATAACTTCAACTACGGGG
 CTTACCATTCCCTGGAAGCTATTTACCACGAGATGGACAACATTGCCGCAGACTTTCTGAC
 CTGGCGAGGAGGGTGAAGATTGGACATTTCGTTTGAAAACCGGCCGATGTATGTACTGAAGTT
 CAGCACTGGGAAAGGCGTGAGGCGGCCGGCCGTTTGGCTGAATGCAGGCATCCATTCCCGAG
 AGTGGATCTCCCAGGCCACTGCAATCTGGACGGCAAGGAAGATTGTATCTGATTACCAGAGG
 GATCCAGCTATCACCTCCATCTTGGAGAAAATGGATATTTTCTTGTTGCCTGTGGCCAATCC
 TGATGGATATGTGTATACTCAAACCTCAAACCGATTATGGAGGAAGACGCGGTCCCGAAATC
 CTGGAAGCTCCTGCATTGGTGCTGACCCAAATAGAACTGGAACGCTAGTTTTGTCAGGAAAG
 GGAGCCAGCGACAACCCCTTGCTCCGAAGTGTACCATGGACCCACGCCAATTCGGAAGTGA
 GGTGAAATCAGTGGTAGATTTTCATCCAAAAACATGGGAATTTCAAGGGCTTCATCGACCTGC
 ACAGCTACTCGCAGCTGCTGATGTATCCATATGGGTACTCAGTCAAAAAGGCCCCAGATGCC
 GAGGAACCTCGACAAGGTGGCGAGGCTTGCGGCCAAAGCTCTGGCTTCTGTGTCGGGCACTGA
 GTACCAAGTGGGTCCCACCTGCACCACTGTCTATCCAGCTAGCGGGAGCAGCATCGACTGGG
 CGTATGACAACGGCATCAAATTTGCATTACATTTGAGTTGAGAGATACCGGGACCTATGGC
 TTCCTCCTGCCAGCTAACAGATCATCCCCACTGCAGAGGAGACGTGGCTGGGGCTGAAGAC
 CATCATGGAGCATGTGCGGGACAACCTCTAC**TAG**GCGATGGCTCTGCTCTGTCTACATTTAT
 TTGTACCCACACGTGCACGCACTGAGGCCATTGTTAAAGGAGCTCTTTCCTACCTGTGTGAG
 TCAGAGCCCTCTGGGTTTGTGGAGCACACAGGCCTGCCCCCTCTCCAGCCAGCTCCCTGGAGT
 CGTGTGTCCTGGCGGTGTCCCTGCAAGAACTGGTTCTGCCAGCCTGCTCAATTTTGGTCCTG
 CTGTTTTTGATGAGCCTTTTGTCTGTTTCTCCTTCCACCCTGCTGGCTGGGCGGCTGCACTC
 AGCATACCCCTTCCTGGGTGGCATGTCTCTCTACCTCATTTTTTAGAACCAGAAACATC
 TGAGATGATTCTCTACCTCATCCACATCTAGCCAAGCCAGTGACCTTGCTCTGGTGGCACT
 GTGGGAGACACCACTTGTCTTTAGGTGGGTCTCAAAGATGATGTAGAATTTCCTTTAATTTT
 TCGCAGTCTTCTGGAAAATATTTTCCCTTTGAGCAGCAAATCTTGTAGGGATATCAGTGAAG
 GTCTCTCCCTCCCTCCTCTCCTGTTTTTTTTTTTTTGGAGACAGAGTTTTGCTCTTGTTGCC
 CAGGCTGGAGTGTGATGGCTCGATCTTGGCTCACCACAACCTCTGCCTCCTGGGTTCAAGCA
 ATTCTCCTGCCTCAGCCTCTTGAGTAGCTTGGTTTATAGGCGCATGCCACCATGCCTGGCTA
 ATTTTGTGTTTTTAGTAGAGACAGGGTTTCTCCATGTTGGTCAGGCTGGTCTCAAACCTCCCA
 ACCTCAGGTGATCTGCCCTCCTTGGCCTCCCAGAGTGCTGGGATTACAGGTGTGAGCCACTG
 TGCCGGGGCCCGTCCCCTCCTTTTTTAGGCCTGAATACAAAGTAGAAGATCACTTTCCTTCAC
 TGTGCTGAGAATTTCTAGATACTACAGTTCTTACTCCTCTCTTCCCTTTGTTATTACAGTGTG
 ACCAGGATGGCGGGAGGGGATCTGTGTCACTGTAGGTACTGTGCCCAGGAAGGCTGGGTGAA
 GTGACCATCTAAATTGCAGGATGGTGAAATTATCCCCATCTGTCCTAATGGGCTTACCTCCT
 CTTTGCCTTTTGAACTCACTTCAAAGATCTAGGCCTCATCTTACAGGTCCTAAATCACTCAT
 CTGGCCTGGATAATCTCACTGCCCTGGCACATTCCCATTTGTGCTGTGGTGTATCCTGTGTT
 TCCTTGTCTGGTTTG
 TCTGTCTATTTTGTATCCTGGACCACAAGTTCCTAAGTAGAGCAAGAATTCATCAACCAGCT
 GCCTCTTGTTCATTTACCTCAGCACGTACCATCTGTCTTTTGTGTTGTGTGTGTGTGTGTG
 TTGTTTTTTGCTTTTACCAAACATGTCTGTAAATCTTAACCTCCTGCCTAGGATTTGTACA
 GCATCTGGTGTGTGCTTATAAGCCAATAAATATTCAATGTGAAAAAAAAAAAAAAAAAAAA

FIGURE 134

MRWILFIGALIGSSICGQEKFFGDQVLRINVRNGDEISKLSQLVNSNNLKLNFWKSPSSFNR
PVDVLVPSVSLQAFKSFLRSQGLEAYVTIEDLQALLDNEDDEMQHNEGQERSNNFNYGAYH
SLEAIYHEMDNIAADFPDLARRVKIGHSFENRPMYVLKFSTGKGVRRPAVWLNAGIHSREWI
SQATAIWTARKIVSDYQRDPAITSILEKMDIFLLPVANPDGYVYTQTQNRLWRKTRSRNPGS
SCIGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKSVVDFIQKHGNFKGFIDLHSY
SQLLMYPYGYSVKKAPDAEELDKVARLAALKALASVSGTEYQVGPTCTTVYPASGSSIDWAYD
NGIKFAFTFELRDTGTYGFLLPANQIIPTAEETWLGLKTIMEHVRDONLY

Signal peptide:

amino acids 1-16

135/249

FIGURE 135

[illegible]

FIGURE 136

MASYLYGVLFVAVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDFAFRLYRRLVLE
TPSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVP
SKDLTLKMGSALFVKKELQLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVKKKTQGK
VVDIIQGLDLLTAMVLVNHIFFKAKWEKPFHLEYTRKNFPFLVGEQVTVQVPMMHQKEQFAF
GVDTELNCFVLQMDYKGDVAFFVLPSKGKMRQLEQALSARTLIKWSHSLQKRWIEVFIPRF
SISASYNLETILPKMGIQNAFDKNADFSGIAKRDSLQVSKATHKAVLDVSEEGTEATAATTT
KFIVRSKDGPSYFTVSFNRTFLMMITNKATDGILFLGKVENPTKS

Signal peptide:

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/001111>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 137

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGC
CTGAGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAG
CCTTTATCTCTTCACCTTCAAGTCCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAG
TCTTGGTACATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAG**ATGA**AAGATGC
AGAAAGGAAATGTTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCC
AATGAGACTAGCACCTCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGC
CACCAACTCTGGGTCCAGTGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCA
GCGTGACCTCCAATGGGGTCAGCATAGTCACCAACTCTGAGTTCATACAACTCCAGTGGG
ATCAGCACAGCCACCAACTCTGAGTTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAA
CTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAC
CCTCCAGTGGGGCCAGCACAGTCACCAACTCTGGGTCCAGTGTGACCTCCAGTGGAGCCAGC
ACTGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGA
GTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAACTCCA
GTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCC
ACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGAGTCCAG
CACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGAACGACCTCCAATGGGG
CTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAAC
TCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACGAC
CTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCTAGCA
CAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACTCTGAG
TCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCTCCAG
TGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCA
CCAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGC
ACAACCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGC
TAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACT
CTGAGTCTAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAACT
TCCAGTGGGGCCAACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAAC
AGCAGCTCTGACTGGAATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAAGGCAA
AGCCTGGTGGGTCCCTGGTGCCGTGGGAAATCTTCCTCATCACCTGGTCTCGGTTGTGGCG
GCCGTGGGGCTCTTTGCTGGGCTCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACAC
CTTTAACACAGCTGTCTACCACCCTCATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAG
GGAATCATGGAGCCCCCACAGGCCAGGTGGAGTCCTAACTGGTTCTGGAGGAGACCAGTA
TCATCGATAGCCATGGAGATGAGCGGGAGGAACAGCGGGCCCT**TGAG**CAGCCCCGGAAGCAAG
TGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCACCCAAGACCTGGTTTCCTTTCATTTCATC
CCAGGAGACCCCTCCCAGCTTTGTTTGAGATCCTGAAAATCTTGAAGAAGGTATTCTCACC
TTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATATTGCTCATTTAGCTAAGAAATAA
ATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTGCCCCGGGGTGGGTATCTAG
CTCTGAGATGAACTCAGTTATAGGAGAAAACCTCCATGCTGGACTCCATCTGGCATTCAAAA
TCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

FIGURE 138

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATI
SGSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSE
SSTPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSS
TTSSGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTT
SNGAGTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSS
GASTATNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGA
NTATNSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEAST
ATNSESSTVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAV
SEAKPGGSLVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGP
GPGGNHGAPHRPRWSPNWFWRPVS SIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 139

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTC
CCAGCAAT**ATG**CATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCC
CTGTCTGGATGGGCGGCCAGCGATGACCCCATTTGAGAAGGTCATTGAAGGGATCAACCGAGG
GCTGAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGGAATCACGC
ATGCCGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGC
AAGGAGTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGAT
CAACCATGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACG
CTGCTGGACAGGCCGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCAC
CAGGCTGGGAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGG
AAAGGAAGTGGAGAAGCTTGGCCAAGGTGCCCACCATGCTGCTGGCCAGGCCGGAAGGAGC
TGCAGAATGCTCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGC
AACCATCAAAGCGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGG
GGCCTCAGTCAACACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCA
TGCCC**TAA**ACTGGCATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTCACATCAGCTGAC
ATGACCTGGAGGGGTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTGTACTG
GGATTTGTGAATAAACTTGATACACCA

FIGURE 140

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675
><subunit 1 of 1, 247 aa, 1 stop
><MW: 25335, pI: 7.00, NX(S/T): 0
MHLARLVGSCSLLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAG
REVEKVFNGLSNMGSHTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHGVNNAAG
QAGKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQN
AHNGVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP
```

Important features of the protein:

Signal peptide:

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

FIGURE 141

CTCCGGGTCCCCAGGGGCTGCGCCGGGCCGGCCTGGCAAGGGGGACGAGTCAGTGGACACTCCAGGAAGAGCGGC
 CCCGCGGGGGGCGATGACCGTGCGCTGACCTGACTCACTCCAGGTCCGGAGGCGGGGGCCCCGGGGCGACTCG
 GGGGCGGACCGCGGGGGCGGAGCTGCCGCCGTGAGTCCGGCCGAGCCACCTGAGCCCCGAGCCGCGGGACACCGTC
 GCTCCTGCTCTCCGAATGCTGCGCACCCGCGATGGGCCGTGAGGAGCTGGCTCGCCGCCCATGGGGCGCGCTGCCG
 CCTCGGCCACCGCTGCTGCTGCTCCTGCTGCTGCTCCTGCTGCGAGCCGCCGCTCCGACCTGGGCGCTCAGC
 CCCCCGATCAGCCTGCCCTCTGGGCTCTGAAGAGCGGCCATTCTCAGATTGGAAGCTGAACACATCTCCAACCTAC
 ACAGCCCTTCTGCTGAGCAGGGATGGCAGGACCTGTACGTGGGTGCTCGAGAGGCCCTCTTTGCACTCAGTAGC
 AACCTCAGCTTCCCTGCCAGGCGGGGAGTACCAGGAGCTGCTTTGGGGTGCAGACGCAGAGAAGAAACAGCAGTGC
 AGCTTCAAGGGCAAGGACCCACAGCGGAGTGTCAAAACTACATCAAGATCCTCCTGCCGCTCAGCGGCAGTCAC
 CTGTTACCTGTGGCACAGCAGCCTTCAAGCCCATGTGTACCTACATCAACATGGAGAATTTACCCCTGGCAAGG
 GACGAGAAGGGGAATGTCTCCTGGAAGATGGCAAGGGCCGTTGTCCCTTCGACCCGAATTTCAAGTCCACTGCC
 CTGGTGGTTGATGGCGAGCTCTACACTGGAACAGTCAGCAGCTTCCAAGGGAATGACCCGGCCATCTCGCGGAGC
 CAAAGCCTTCGCCCCACCAAGACCGAGAGCTCCCTCAACTGGCTGCAAGACCCAGCTTTTGTGGCCTCAGCCTAC
 ATTCTGAGAGCCTGGGCAGCTTGAAGGCGATGATGACAAGATCTACTTTTTCTTCAGCGAGACTGGCCAGGAA
 TTTGAGTTCTTTGAGAACACCAATTTGTCTCCCGCATTTGCCGCATCTGCAAGGGCGATGAGGGTGGAGAGCGGGTG
 CTACAGCAGCGCTGGACCTCCTTCTCAAGGCCAGCTGCTGTGCTCACGGCCCGACGATGGCTTCCCTTCAAC
 GTGCTGCAGGATGTCTTCAGCTGAGCCCCAGCCCCAGGACTGGCGTGACACCTTTTCTATGGGGTCTTCACT
 TCCCAGTGGCACAGGGGAACTACAGAAGGCTCTGCCGTCTGTGTCTTCAATGAAGGATGTGCAGAGAGTCTTC
 AGCGGCTCTACAAGGAGGTGAACCGTGAGACACAGCAGTGGTACACCGTGACCCACCCGGTGCCACACCCCGG
 CCTGGAGCGTGCAATCAACAAGTGGCCGGGAAAGGAAGTCAACTCATCCTGCAGCTCCCAGACCGCGTGCTG
 AACTTCTCAAGGACCACTTCTGATGGACGGGCAGGTCCGAAGCCGATGCTGCTGCTGCAGCCCCAGGCTCGC
 TACCAGCGGTGGCTGTACACCGCTCCCTGGCCTGCACCACACCTACGATGTCTCTTCTGGGCACTGGTGAC
 GGCCGGTCCACAAGGCAGTGAGCGTGGGCCCCGGGTGCACATCATTGAGGAGCTGCAGATCTTCTCATCGGGA
 CAGCCCGTGAGAATCTGCTCCTGGACACCCACAGGGGGCTGCTGTATGCGGCCTCACACTCGGGCGTAGTCCAG
 GTGCCCATGGCCAATGCAGCCTGTACCGGAGCTGTGGGACTGCCTCCTCGCCCGGGACCCCTACTGTGCTTGG
 AGCGGCTCCAGCTGCAAGCACGTGAGCCTCTACAGCCTCAGCTGGCCACCAGGCGGTGGATCCAGGACATCGAG
 GGAGCCAGCGCAAGGACCTTTGACGCGCTTTCGGTTGTGTCCCGCTCTTTTGTACCAACAGGGGAGAAGCCA
 TGTGAGCAAGTCCAGTTCCAGCCCAACACAGTGAACACTTTGGCCTGCCCGCTCCTCTCCAACCTGGCGAGCCGA
 CTCTGGCTACGCAACGGGGCCCCCGTCAATGCCTCGGCCTCCTGCCACGTGCTACCCACTGGGGACCTGCTGCTG
 GTGGGCAACCAACAGCTGGGGGAGTTCCAGTGTGCTGGTCACTAGAGGAGGGCTTCCAGCAGCTGGTAGCCAGCTAC
 TGCCAGAGGTGGTGGAGGACGGGGTGGCAGACCAACAGATGAGGGTGGCAGTGTACCCGTCAATATCAGCACA
 TGCCGTGTGAGTGACACAGCTGGTGGCAAGGCCAGCTGGGGTGCAGACAGTCTACTGGAAGGAGTCTCTGGTG
 ATGTGCACGCTCTTTGTGCTGGCCGTGCTGCTCCAGTTTATTTCTTGCTCTACCGGCACCGGAACAGCATGAAA
 GTCTTCTGAAGCAGGGGGAATGTGCCAGCGTGACCCCAAGACCTGCCCTGTGGTGCTGCCCCCTGAGACCCGC
 CCACTCAACGGCCCTAGGGCCCCCTAGCACCCCGCTCGATCACCGAGGGTACCAGTCCCTGTGAGACAGCCCCCG
 GGGGCCCGAGTCTTCACTGAGTCAGAGAAGAGGCCACTCAGCATCCAAGACAGCTTCGTGGAGGTATCCCCAGTG
 TGCCCCCGGGCCCCGGGTCCGCCTTGGCTCGGAGATCCGTGACTCTGTGGTGTGAGAGCTGACTTCCAGAGGACGC
 TGCCCTGGCTTCAGGGGCTGTGAATGCTCGGAGAGGGTCACTGGACCTCCCTCCGCTCTGCTCTTCGTGGAAC
 ACGACCGTGGTGCCCGGCCCTTGGGAGCCTTGGAGCCAGCTGGCCTGCTGCTCTCCAGTCAAGTAGCGAAGCTCC
 TACCACCCAGACACCCAAACAGCCGTGGCCCCAGAGGTCCCTGGCCAAATATGGGGGCTGCCTAGGTTGGTGGA
 CAGTGCTCCTTATGTAAACTGAGCCCTTTGTTTAAAAAACAATTCCAAATGTGAAACTAGAATGAGAGGGAAGAG
 ATAGCATGGCATGCAGCACACAGGCTGCTCCAGTTTATGGCCTCCAGGGGTGCTGGGGATGCATCCAAAGTGG
 TTGCTGTGAGACAGAGTTGGAACCCCTACCAACTGGCCTCTTCACTTCCACATTATCCCCTGCCACCGGCTGC
 CCTGTCTCACTGCAGATTGAGGACAGCTTGGGCTGCTGCGTTCTGCCTTGCCAGTCAGCCGAGGATGTAGTTG
 TTGCTGCCGTGCTCCACACCTCAGGGACCAGAGGGCTAGGTTGGCACTGCGGCCCTCACCAGGTCTGGGCTC
 GGACCCAACTCCTGGACCTTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACAGCGGAGCTCAGGAGAGA
 TTTCTGTGACAATGTACGCCTTTCCCTCAGAATTCAGGGAAGAGACTGTGCGCTGCCCTTCTCCGTTGTTGCGTGA
 GAACCCGTGTGCCCTTTCCACCATATCCACCCTCGCTCCATCTTTGAACTCAAACACGAGGAACCTAACTGCACC
 CTGGTCTCTCCCGAGTCCCGAGTTTACCCTCCATCCCTCACCTTCTTCACTCTAAGGGATATCAACACTGCCC
 AGCACAGGGGCCCTGAATTTATGTGGTTTTTATACATTTTTTAATAAGATGCACTTTATGTATTTTTTAATAAA
 GTCTGAAGAATTACTGTTTTAAAAA

FIGURE 142

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67962

><subunit 1 of 1, 837 aa, 1 stop

><MW: 92750, pI: 7.04, NX(S/T): 6

MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLLLLQPPPPTWALSPRISLPLGSEERPFLRF
EAEHISNYTALLLSRDGRTLYVGAREALFALSSNLSFLPGGEYQELLWGADAEKKQQCSFKG
KDPQRDCQNYIKILLPLSGSHLFTCGTAAFSMCTYINMENFTLARDEKGNVLLEDGKGRCP
FDPNFKSTALVVDGELYTGTVSSSQGNDPAISRSQSLRPTKTESSLNWLQDPAFVASAYIPE
SLGSLQGDDDKIYFFFSETGQFEFFENTIVSRIARICKGDEGGERVLQQRWTSFLKAQLLC
SRPDDGFPPFNVLQDVFTLSPSPQDWRDTLFYGVFTSQWHRGTTEGSAVCVFTMKDVQRVFSG
LYKEVNRETQQWYTVTHPVPTPRPGACITNSARERKINSSLQLPDRVLNFLKDHFLMDGQVR
SRMLLLQPQARYQRVAVHRVPGLHHTYDVLFLGTGDGRLHKAVSVGPRVHIIIEELQIFSSGQ
PVQNLLLDTHRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSLEYQP
QLATRPWIQDIEGASAKDLCSASSVVSPSFVPTGEKPCEQVQFQPNVTNLTACPLLSNLATR
LWLRNGAPVNASASCHVLPTGDLLLVTGQQLGEFQCWSLEEGFQQLVASYCPEVVEDGVADQ
TDEGGSVPVPIISTSRVSAPAGGKASWGADRSYWKEFLVMCTLFVLAVLLPVLFLLYRHRNSM
KVFLKQGECAVHPKTCPVVLPPETRPLNGLGPPSTPLDHRGYQSLSDSPPGARVFTSEKR
PLSIQDSFVEVSPVCPRPRVRLGSEIRDSVV

Transmembrane domains:

amino acids 23-46 (type II), 718-738

FIGURE 143A

CTAAGCCGGAGGATGTGCAGCTGCGGCGGGCGCGCCGGCTACGAAGAGGACGGGGACAGGCGCCGTGCGAACCGA
 GCCCAGCCAGCCGGAGGACGCGGGCAGGGCGGGACGGGAGCCCCGGACTCGTCTGCCGCGCCGCTCGTCCGCGTCCG
 TGCCGGCCCCGCGTCCCCGCGCGAGCGGGAGGAGCCGCCGCCACCTCGCGCCCGAGCCCGCGCTAGCGCGCGC
 CGGGCATGGTCCCCTCTTAAAGGCGCAGGCCGCGCGGGCGGGGGCGGGTGTGCGGAACAAAGCGCCGGCGCGGGG
 CCTGCGGGCGGCTCGGGGGCCGCGATGGGCGCGCGGGCCCCGCGCGCGCGCGCGCTGCCCGGGCCGGGCCTCG
 CGGCGCTAGGGCGGGCTGGCCTCCGTGGGCGGGGGCAGCGGGCTGAGGGCGCGCGGAGCCTGCGGCGGGCGGGCGG
 GCGGGCGGGCGGGCGGGCGGGAGCGGGCGGGGCGGGGATGCGGGCGCGCGGGCGGGCGGGCGGGCTGGCTCAGCGTGC
 TGCTCGGGCTCGTCTGGGCTTCGTGCTGGCCTCGCGGCTCGTCTGCCCCGGGCTTCGAGCTGAAGCGAGCGG
 GCCCAGGGCGCGCGCCAGCCCCGAGGGCTGCCGCTCGGGCAGGCGGGCGGCTTCCAGGCCGGCGGGCGCGCG
 GCGATGCGCGCGGGGCGCAGCTCTGGCCGCCCGGCTCGGACCCAGATGGCGGGCGCGCGACAGGAACCTTCTCT
 TCGTGGGAGTCATGACCGCCAGAAATACCTGCAGACTCGGGCCGTGGCCGCTACAGAACATGGTCCAAGACAA
 TTCTTGGGAAAGTTCACTTCTTCAAGTGAGGGTTCTGACACATCTGTACCAATTCCAGTAGTGCCACTACGGG
 TTGTGGACGACTCCTACCGCCCCAGAAGAAGTCTTTCATGATGCTCAAGTACATGCACGACCACTACTTGGACA
 AGTATGAATGGTTTATGAGAGCAGATGATGACGTGATACATCAAAGGAGACCGTCTGGAGAATTCTTGGAGGTT
 TGAACAGCAGCGAGCCCCCTCTTCTTGGGCGAGACAGGCGCTGGGCGCCACCGGAAGAAATGGGAAAACCTGGCCCTGG
 AGCCTGGTGAGAACTTCTGCATGGGGGGGCGCTGGCGTGATCATGAGCCGGGAGGTGCTTCGGAGAATGGTGGCG
 ACATTGGCAAGTGTCTCCGGGAGATGTACACCACCCATGAGGACGTGGAGGTGGGAAGGTGTGTCCGGAGGTTTG
 CAGGGGTGCAGTGTGTCTGGTCTTATGAGATGCGGCAGCTTTTTTATGAGAATTACGAGCAGAACAAAAAGGGGT
 ACATTAGAGATCTCCATAACAGTAAATTCACCAAGCTATCACATTACCCCCAACAAAAACCCACCTTACCAGT
 ACAGGCTCCACAGCTACATGCTGAGCCCGCAAGATATCCGAGCTCCGCCATCGCACAAATACAGCTGCACCGCGAAA
 TTGTCTGATGAGCAAATACAGCAACACAGAAATTCATAAAGAGGACCTCCAGCTGGGAATCCCTCCCTCCTTCA
 TGAGGTTTCAGCCCCGCCAGCGAGAGGAGATTCTGGAATGGGAGTTTCTGACTGGAAAATACTTGTATTTCGGCAG
 TTGACGGCCAGCCCCCTCGAAGAGGAATGGAATCCGCCAGAGGGAAGCCTTGGACGACATTGTCTATGCAGGTCA
 TGGAGATGATCAATGCCAACGCCAAGACCAGAGGGCGCATATTGACTTCAAAGAGATCCAGTACGGCTACCGCC
 GGGTGAACCCCATGTATGGGGCTGAGTACATCCTGGACCTGCTGCTTCTGTACAAAAGCACAAAGGGGAAGAAAA
 TGACGGTCCCTGTGAGGAGGCACGCGTATTTACAGCAGACTTTCAGCAAAATCCAGTTTGTGGAGCATGAGGAGC
 TGGATGCACAAGAGTTGGCCAAGAGAATCAATCAGGAATCTGGATCCTTGTCTTCTCTCAAACTCCCTGAAGA
 AGCTCGTCCCCTTTCAGCTCCCTGGGTGGAAGAGTGAGCACAAAGAACCCAAAGATAAAAAAGATAAACTGTA
 TTCTTTGTCTGGGCGTTTTGACATGTTTTGTGAGATTTATGGGAACTTTGAGAAGACGTGTCTTATCCCCAATC
 AGAAGCTCAAGCTCGTGGTTCTGCTTTTCAATTCTGACTCCAACCCTGACAAAGGCCAAACAGTTGAAGTATGA
 GAGATTACCGCATTAAGTACCCTAAAGCCGACATGCAGATTTTGCCTGTGTCTGGAGAGTTTTCAAGAGCCCTGG
 CCCTGGAAGTAGGATCCTCCAGTTTAAACAATGAATCTTGTCTTCTTCTGCGACGTGACCTCGTGTTTACTA
 CAGAATTCCTTCAGCGATGTGAGCAAATACAGTTCTGGGCCAACAAATATATTTTCCAATCATCTACAGCAGT
 ATGACCCAAAGATTGTTTATAGTGGGAAAGTTCCAGTGACAACCATTTTGCCTTTACTCAGAAAACCTGGCTTCT
 GGAGAACTATGGGTTTGGCATCACGTGATTTTATAAGGGAGATCTTGTCCGAGTGGGTGGCTTTGATGTTTCCA
 TCCAAGGCTGGGGGCTGGAGGATGTGGACCTTTTCAACAAGGTTGTCCAGGCAGGTTTGAAGACGTTTAGGAGCC
 AGGAAGTAGGAGTAGTCCAGCTCCACCATCCTGTCTTTTGTGATCCCAATCTTGACCCCAAACAGTACAAAATGT
 GCTTGGGGTCCAAAGCATCGACCTATGGGTCCACCCAGCAGTGGCTGAGATGTGGCTGGAAAAAAGATGATCCAA
 GTTACAGTAAAAGCAGCAATAATAATGGCTCAGTGAGGACAGCCTAAATGTCCAGCTTTGCTGGAAAAAGACGTTTT
 TAATTATCTAATTTATTTTTCAAAAATTTTTTGTATGATCAGTTTTTGAAGTCCGTATACAAGGATATATTTTAC
 AAGTGGTTTTCTTACATAGGACTCCTTTAAGATTGAGCTTTCTGAACAAGAAGGTGATCAGTGTGCTTTGAA
 CACATCTTCTGCTGAACATTATGTAGCAGACCTGCTTAACTTTGACTTGAAATGTACCTGATGAACAAAACCTT
 TTTAAAAAATGTTTTCTTTTGGAGCCCTTGTCTCAGTCTTATGGCAGAAAACGTGAACATTCTGCAAAGTAT
 TATTGTAACAAAACACTGTAACCTCTGGTAAATGTTCTGTTGTGATTGTTAACATTCCACAGATTCTACCTTTGT
 GTTTTGTTTTTTTTTTTTTTACAATTGTTTTAAAGCCATTTTCATGTTCCAGTTGTAAGATAAGGAAATGTGATAATA
 GCTGTTTCATCATTGTCTTTCAGGAGAGCTTTCAGAGTTGATCATTCTCTCATGGTACTCTGCTCAGCATGGC
 CACGTAGGTTTTTTGTTTGTGTTTTGTTTCTTTTTTGGAGACGGAGTCTCACTCTGTTACCCAGGCTGGAATG
 CAGTGGCGCAATCTTGGCTCACTTTAACCTCCACTTCCCTGGTTCAGCAATTCCCCTGCCTTTGCCTCCCGAGT
 AGCTGGGATTACAGGCACACACCACCGCCAGNTAGTTTTTTTGTATTTTTAGTAGAGACGGGTTTTACCATT
 GCAAGCCCAGCTGGCCACGTAGGTTTTAAAGCAAGGGGCGTTGAAGAAGGCACAGTGAGGTATGTGGCTGTTCTCG
 TGGTAGTTTCATTGCGCCTAAATAGACCTGGCATTAATTTCAAGAAGGATTTGGCATTTTCTCTTCTTGACCCTT
 CTCTTTAAAGGGTAAATATTAATGTTTGAATGACAAAGATGAATTATTACAATAAATCTGATGTACACAGACT
 GAAACATACACACATACACCCTAATCAAAACGTTGGGGAAAAATGTATTTGGTTTTGTTTCTTTTCATCCTGTCTG
 TGTATGTGGTGGAGATGGTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
 TTTATTTAATATCTGTTTGTTCAGAGCTCTGCCATTTCTTGAGTACCTGTTAGTTAGTATTATTTATGTGTATCGG
 GAGTGTGTTTAGTCTGTTTTATTTGCAGTAAACCGATCTCCAAGATTTTCTTTTGGAAACGCTTTTTCCCTCC

TTAATTTTTATATTCTTACTGTTTTACTAAATATTAAGTGTTCTTTGACAATTTTGGTGCTCATGTGTTTTGGG
GACAAAAGTGAAATGAATCTGTCTATTAACGAGAAAGTAAATTTCTCAGATCAAATGTGCCTTAATAAATTTGTT
TTCATTTTAGATTTCAAACAGTGATAGACTTGCCATTTTAAATACACGTCATTTGGAGGGGCTGCGTATTTGTAAATAG
CCTGATGCTCATTTTGGAAAAATAAACCAAGTGAACAATATTTTCTATTGTACATTTTGC AACCATTTTGTCTCATTT
ATTCCTGTTTTAGCTGAAGAATTGTATTACATTTGGAGAGTAAAAAACTTAAACACGAAAAAA

FIGURE 144

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68836

><subunit 1 of 1, 802 aa, 1 stop

><MW: 91812, pI: 9.52, NX(S/T): 3

MAARGRRRAWLSVLLGLVLGFLASRLVLPRASELKRAGPRRRASPEGCRSGQAAASQAGGAR
GDARGAQLWPPGSDPDGGPRDRNFLVGVMTAQKYLQTRAVAAYRTWSKTIPGKVQFFSSEG
SDTSVPIPVVPLRGVDDSYPPQKKSEFMLKYMHDHYLDKYEWFMRADDDVYIKGDRLENFLR
SLNSSEPLFLGQTGLGTTEEMGKLALEPGENFCMGGPGVIMSREVLRRMVPHIGKCLREMYT
THEDVEVGRCVRRFAGVQCVWSYEMRQLFYENYEQNKKGYIRDLHNSKIHQAITLHPNKNPP
YQYRLHSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGIPPSFMRFPQRQREE
ILEWEFLTGKYLYSAVDGQPPRRGMDSAQREALDDIVMQVMEMINANAKTRGRIIDFKEIQY
GYRRVNPYGAEYILDLLLLYKKHKGKKMTVPVRRHAYLQQTFSKIQFVEHEELDAQELAKR
INQESGSLSFSLNSLKKLVFPQLPGSKSEHKEPKDKKINILIPLSGRFDMFVRFMGNFEKTC
LIPNQNVKLVLLFNSDSNPDKAKQVELMRDYRIKYPKADMQILPVSGEFSRALALEVGSSQ
FNNESLLFFCDVDLVFTTEFLQRCRANTVLGQQIYFPIIFSQYDPKIVYSGKVPSDNHFAFT
QKTGEWRNYGFGITCIYKGD LVRVGGFDVSIQGWGLEDDVLFNKVVQAGLKTFRSQEVGVVH
VHHPVFCDPNLDPKQYKMCLGSKASTYGSTQQLAEMWLEKNDPSYSKSSNNNGSVRTA

Signal peptide:

amino acids 1-23

FIGURE 145

GGACAACCGTTGCTGGGTGTCCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCC
 TTTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGCT
 ACGGACGACGCCT**ATGA**AGCCCTTAGTCCTTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTG
 TGCCGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAACTTGAATCATTATATA
 CAAGTTTTAGAGAACCTAGTACGAAGTGTTCCCTCTGGGGAGCCAGGTCGTGAGAAAAAATC
 TAACTCTCCAAAACATGTTTATTCTATAGCATCAAAGGGATCAAATTTAAGGAGCTAGTTA
 CACATGGAGACGCTTCAACTGAGAATGATGTTTTAACCAATCCTATCAGTGAAGAACTACA
 ACTTTCCTTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAAACACACGGAAAGTACCCC
 ATTCTGGTCGATCAAACCAAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTG
 AAAATGAAGAGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAACTGAGGCACCAAGAATG
 TTGCCAGTTGTTACTGAATCATCTACAAGTCCATATGTTACCTCATAACAAGTCACCTGTCAC
 CACTTTAGATAAGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCCTCAGCTCT
 CAGGTGAAACTGCGATAGAAAAACCCGAAGAGTTTGGAAGCACCCAGAGAGTTGGAATAAT
 GATGACATTTTGAAAAAAATTTTAGATATTAATTCACAAGTGCAACAGGCACTTCTTAGTGA
 CACCAGCAACCCAGCATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCC
 TTGCTCTAGCAGCAGCAGCAGAACATAAATTAAAAACAATGTATAAGTCCCAGTTATTGCCA
 GTAGGACGAACAAGTAATAAAATTGATGACATCGAAACTGTTATTAACATGCTGTGTAATTC
 TAGATCTAAACTCTATGAATATTTAGATATTAAATGTGTTCCACCAGAGATGAGAGAAAAAG
 CTGCTACAGTATTCAATACATTAAAAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTA
 AAAGTTTAT**TAA**ACAATAATATAAAAAATTTTAAACCTACTTGATATTCCATAACAAAGCTGA
 TTTAAGCAAACCTGCATTTTTTTCACAGGAGAAATAATCATATTCGTAATTTCAAAGTTGTAT
 AAAAATATTTTCTATTGTAGTTCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACA
 ATTTTCATATGCACTAAAAACCTAATTTAAATAAAATTTTGGTTCAGGAAAAA

FIGURE 146

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68864
><subunit 1 of 1, 350 aa, 1 stop
><MW: 39003, pI: 5.59, NX(S/T): 1
MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRVPSGEPGREKKSNSPK
HVYSIASKGSKFKELVTHGDASTENDVLNPISEETTTFTGGFTPEIGKKKHTESTPFWSI
KPNNVSIVLHAEOPYIENEEPEPEPEPAAKQTEAPRMLPVVTESSTSPYVTSYKSPVTTLDK
STGIEISTESEDVPQLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNP
AYREDIEASKDHLKRSLALAAAAEHKLKTMYSQLLPVGRTSNKIDDIETVINMLCNSRSKL
YEYLDIKCVPEMREKAATVFNTLKNMCRSRRVTALLKVY
```

Signal peptide:

amino acids 1-19

1001 999 998 997 996 995 994 993 992 991 990 989 988 987 986 985 984 983 982 981 980 979 978 977 976 975 974 973 972 971 970 969 968 967 966 965 964 963 962 961 960 959 958 957 956 955 954 953 952 951 950 949 948 947 946 945 944 943 942 941 940 939 938 937 936 935 934 933 932 931 930 929 928 927 926 925 924 923 922 921 920 919 918 917 916 915 914 913 912 911 910 909 908 907 906 905 904 903 902 901 900 899 898 897 896 895 894 893 892 891 890 889 888 887 886 885 884 883 882 881 880 879 878 877 876 875 874 873 872 871 870 869 868 867 866 865 864 863 862 861 860 859 858 857 856 855 854 853 852 851 850 849 848 847 846 845 844 843 842 841 840 839 838 837 836 835 834 833 832 831 830 829 828 827 826 825 824 823 822 821 820 819 818 817 816 815 814 813 812 811 810 809 808 807 806 805 804 803 802 801 800 799 798 797 796 795 794 793 792 791 790 789 788 787 786 785 784 783 782 781 780 779 778 777 776 775 774 773 772 771 770 769 768 767 766 765 764 763 762 761 760 759 758 757 756 755 754 753 752 751 750 749 748 747 746 745 744 743 742 741 740 739 738 737 736 735 734 733 732 731 730 729 728 727 726 725 724 723 722 721 720 719 718 717 716 715 714 713 712 711 710 709 708 707 706 705 704 703 702 701 700 699 698 697 696 695 694 693 692 691 690 689 688 687 686 685 684 683 682 681 680 679 678 677 676 675 674 673 672 671 670 669 668 667 666 665 664 663 662 661 660 659 658 657 656 655 654 653 652 651 650 649 648 647 646 645 644 643 642 641 640 639 638 637 636 635 634 633 632 631 630 629 628 627 626 625 624 623 622 621 620 619 618 617 616 615 614 613 612 611 610 609 608 607 606 605 604 603 602 601 600 599 598 597 596 595 594 593 592 591 590 589 588 587 586 585 584 583 582 581 580 579 578 577 576 575 574 573 572 571 570 569 568 567 566 565 564 563 562 561 560 559 558 557 556 555 554 553 552 551 550 549 548 547 546 545 544 543 542 541 540 539 538 537 536 535 534 533 532 531 530 529 528 527 526 525 524 523 522 521 520 519 518 517 516 515 514 513 512 511 510 509 508 507 506 505 504 503 502 501 500 499 498 497 496 495 494 493 492 491 490 489 488 487 486 485 484 483 482 481 480 479 478 477 476 475 474 473 472 471 470 469 468 467 466 465 464 463 462 461 460 459 458 457 456 455 454 453 452 451 450 449 448 447 446 445 444 443 442 441 440 439 438 437 436 435 434 433 432 431 430 429 428 427 426 425 424 423 422 421 420 419 418 417 416 415 414 413 412 411 410 409 408 407 406 405 404 403 402 401 400 399 398 397 396 395 394 393 392 391 390 389 388 387 386 385 384 383 382 381 380 379 378 377 376 375 374 373 372 371 370 369 368 367 366 365 364 363 362 361 360 359 358 357 356 355 354 353 352 351 350 349 348 347 346 345 344 343 342 341 340 339 338 337 336 335 334 333 332 331 330 329 328 327 326 325 324 323 322 321 320 319 318 317 316 315 314 313 312 311 310 309 308 307 306 305 304 303 302 301 300 299 298 297 296 295 294 293 292 291 290 289 288 287 286 285 284 283 282 281 280 279 278 277 276 275 274 273 272 271 270 269 268 267 266 265 264 263 262 261 260 259 258 257 256 255 254 253 252 251 250 249 248 247 246 245 244 243 242 241 240 239 238 237 236 235 234 233 232 231 230 229 228 227 226 225 224 223 222 221 220 219 218 217 216 215 214 213 212 211 210 209 208 207 206 205 204 203 202 201 200 199 198 197 196 195 194 193 192 191 190 189 188 187 186 185 184 183 182 181 180 179 178 177 176 175 174 173 172 171 170 169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154 153 152 151 150 149 148 147 146 145 144 143 142 141 140 139 138 137 136 135 134 133 132 131 130 129 128 127 126 125 124 123 122 121 120 119 118 117 116 115 114 113 112 111 110 109 108 107 106 105 104 103 102 101 100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

FIGURE 147

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTG
ACCTCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACA
GGAATATCC**ATGG**CCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACA
GTGGCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCT
CCTGCTCCCTCTTTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAG
TTCCATGCTGTGGTCCACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACA
GTATCGAGGGGAGAACTGAGTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGC
TAAAAAACATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTTCTCAGTTCCCAGATTTAC
GATGAGGAGGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCAT
CGTGGGATATGTTGACGGAGGTATCCAGTTACTCTGCCTGTCTCAGGCTGGTTCCCCCAGC
CCACAGCCAAGTGGAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCA
GATGGGTACAGCCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAAATGCTGGGAGCAT
ATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAG
AGACGTTTTTCCAGCCCTCACCTTGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGT
GCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTCAAATCCAAAGGGAAAATCCA
GGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCCGAAACACG
CAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAA
ACTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAA
GAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGAC
AAAATGTAGGGTGGTATGTGGGAGTGTGTGCGGATGACGTAGACAGGGGGAAGAACAATGTG
ACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTTAC
ATTCAATCCCCATTTTATCAGCCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCC
TGGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATAACC
CTGCTGACATGTCAGTTTGAAGGCTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGA
GGAAAAGGGGACTCCCATATTCATATGTCCAGTGTCTTGGGGA**TGAG**ACAGAGAAGACCCTG
CTTAAAGGGCCCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGC
CCCAGCTTCCTCTCCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGC
TGAGGTTCCTCTGCCCTGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGAT
TGGCCTGACCCTGTGGGAGTCAGAAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCA
CATTAGGTTTGTGTTTGTGAAAACCTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCC
CAGGCTCCTCATTTGCTAGTCACGGACAGTGATTCCCTGCCTCACAGGTGAAGATTAAAGAGA
CAACGAATGTGAATCATGCTTGCAGGTTTGAAGGCACAGTGTTTGCTAATGATGTGTTTTTA
TATTATACATTTTCCCACCATAAACTCTGTTTGCTTATTCCACATTAATTTACTTTTCTCTA
TACCAAATCACCCATGGAATAGTTATTGAACACCTGCTTTGTGAGGCTCAAAGAATAAAGAG
GAGGTAGGATTTTCACTGATTCTATAAGCCAGCATTACCTGATACCAAACAGGCAAAG
AAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCCCTCATTAACACAGACACAAAAA
TTCTAAATAAAATTTTAAACAAATTAACTAAACAATATATTTAAAGATGATATATAACTACT
CAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAAATATTTAAATATCAACCAGTGTAATT
CAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 148

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68866

><subunit 1 of 1, 466 aa, 1 stop

><MW: 52279, pI: 6.16, NX(S/T): 2

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHA
VVHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEE
ATWELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGY
SLYDVEISIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQSPWRLASILLGLLCGALC
GVVMGMIIVFFKSKGKIQAELDWRKKGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVT
HRKAPQEVPHSEKRFTRKSVVASQGFQAGRHYWEVDVGQNVGWYGVCRDDVDRGKNNVTLS
PNNGYWVLRLLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLT
CQFEGLLRPYIQHAMYDEEKGTPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 149

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGGCTA
 GGAAAAGAGTTTGTGGGAACCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTC
 CTGGCAGTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAGAAGACCTACAA
 TTACTATAGCACATTGTCAATTACAACTGACAACTATATGCTGAGTTTGGCAGAGAGGCTT
 CTAACAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAA
 TCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGG
 AGTGTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAG
 ATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTA
 GATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCA
 TTGCTGCGGAACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAG
 AAGTAGAAGAGGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGT
 GGAGCAACCTTAATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTTTACAACATATAA
 GAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACAATAAACCTTCGAAAATGAAACGGG
 GTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCT
 CTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGA
 TGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAA
 ATGATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAAC
 TGCAATGAACCTCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTT
 AGAAGGAAAAACAGATGCATGCCAGGGTGAAGTCTGGAGGACCACTGGTTAGTTCAGATGCTA
 GAGATATCTGGTACCTTGCTGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCAACAAG
 CCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTGGATTACTTCAAAAAGTGGTATCTA
AGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATT
 TTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAAC
 TGTTTGCTTGATGCATGTATTTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTG
 CCAGATCAACTCTGTATCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATA
 ATACAATATTACATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTTGTGAGAATTTTGAC
 TTGTTGACATAAATTTGTAATGCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTC
 AGCTCCTCTCATTTTCAGCAAATATCCATTTTCAAGGTGCAGAACAAGGAGTGAAAGAAAATA
 TAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGT
 GGAATATTAGAAATGATCATATTTCATTATGAAAGGTCAAGCAAAGACAGCAGAATACCAATC
 ACTTCATCATTTAGGAAGTATGGGAAGTAAAGTTAAGGAAGTCCAGAAAGAAGCCAAGATATA
 TCCTTATTTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTG
 ACCTATAATAATTATACAACTTCATGCAATGTACTTGTTCTAAGCAAATTAAAGCAAATAT
 TTATTTAACATTGTTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 150

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68871
><subunit 1 of 1, 423 aa, 1 stop
><MW: 47696, pI: 8.96, NX(S/T): 3
MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTD
KLYAEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICR
FHSTEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTL
GQSLRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFG
VTIKPSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSVPYPYTNVHRVCLPDASYEFQPGDV
MFVTGFGALKNDGYSQNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGD
SGGPLVSSDARDIWYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI
```

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 151

GTCTGAAGGTTATAAAAGCTTCCAGCCAAACGGCATTGAAGTTGAAGATACAACCTGACAGCA
CAGCCTGAGATCTTGGGGATCCCTCAGCCTAACACCCACAGACGTCAGCTGGTGGATTCCCG
CTGCATCAAGGCCTACCCACTGTCTCC**ATG**CTGGGCTCTCCCTGCCTTCTGTGGCTCCTGGC
CGTGACCTTCTTGTTTCCCAGAGCTCAGCCCTTGGCCCCCTCAAGACTTTGAAGAAGAGGAGG
CAGATGAGACTGAGACGGCGTGGCCGCCTTTGCCGGCTGTCCCCTGCGACTACGACCACTGC
CGACACCTGCAGGTGCCCTGCAAGGAGCTACAGAGGGTCGGGGCCGGCGGCCTGCCTGTGCCC
AGGACTCTCCAGCCCCGCCAGCCGCCGACCCGCCGCGCATGGGAGAAGTGCGCATTGCGG
CCGAAGAGGGGCCGCGCAGTGGTCCACTGGTGTGCCCCCTTCTCCCCGGTCCTCCACTACTGG
CTGCTGCTTTGGGACGGCAGCGAGGCTGCGCAGAAGGGGGCCCCCGCTGAACGCTACGGTCCG
CAGAGCCGAAGTGAAGGGGGCTGAAGCCAGGGGGGCATTTATGTCGTTTGCCTAGTGGCCGCTA
ACGAGGCCGGGGCAAGCCGCGTGCCCCAGGCTGGAGGAGAGGGCCTCGAGGGGGCCGACATC
CCTGCCTTCGGGGCCTTGACAGCCGCCTTGCGGTGCCGCCCAACCCCCGCACTCTGGTCCACGC
GGCCGTGGGGGTGGGCACGGCCCTGGCCCTGCTAAGCTGTGCCGCCCTGGTGTGGCACTTCT
GCCTGCGCGATCGCTGGGGCTGCCCGCGCCGAGCCGCCGCCCGAGCCGCAGGGGCGCTC**TGA**
AAGGGGCCTGGGGGCATCTCGGGCACAGACAGCCCCACCTGGGGCGCTCAGCCTGGCCCCCG
GGAAAGAGGAAAACCCGCTGCCTCCAGGGAGGGCTGGACGGCGAGCTGGGAGCCAGCCCCAG
GCTCCAGGGCCACGGCGGAGTCATGGTTCTCAGGACTGAGCGCTTGTTTAGGTCCGGTACTT
GGCGCTTTGTTTCCTGGCTGAGGTCTGGGAAGGAATAGAAAGGGGGCCCCCAATTTTTTTTTTA
AGCGGCCAGATAATAAATAATGTAACCTTTGCGGTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 152

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68874
><subunit 1 of 1, 238 aa, 1 stop
><MW: 25262, pI: 6.44, NX(S/T): 1
MLGSPCLLWLLAVTFLVPRAQPLAPQDFEEEEADETETAWPPLPAVPCDYDHCRHLQVPCKE
LQRVGPAAACLCPLSSPAQPPDPPRMGEVRIAAEEGRAVVHWCAFPSPVLHYWLLLWDGSEA
AQKGPPLNATVRRRAELKGLKPGGIYVVCVVAANEAGASRVPQAGGEGLEGADIPAFGPCSRL
AVPPNPRTLVAHAVGVGTALALLSCAALVWHFCLRDRWGCPRRRAARAAGAL
```

Important features of the protein:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 194-220

N-glycosylation site.

amino acids 132-135

FIGURE 153

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCC
 CTGCCCCGATGAGCCCCCGCGTGCCTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGG
 CCCAGCGCCGACGATCGCTGCCGTTTTGCCCCTGGGAGTAGGATGTGGTGAAAGGATGGGGC
 TTCTCCCTTACGGGGCTCACAA**ATG**GCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCT
 CTACGCCCTCAATCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGA
 TGAGGGACTACCTAAATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTC
 ATTTTGACTTACTTTCCTGTGGTTCATCCGGTCATGATTGCTGTTTGCTGTTTCTTATCAT
 TGTGGGGATGTTAGGATATTGTGGAACGGTGAAAAGAAATCTGTTGCTTCTTGTCATGGTACT
 TTGGAAGTTTGCTTGTCATTTTCTGTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAG
 GAACTTATGGTTCAGTACAATGGTCAGATATGGTCACCTTTGAAAGCCAGGATGACAAATTA
 TGGATTACCTAGATATCGGTGGCTTACTCATGCTTGGAATTTTTTTTCAGAGAGAGTTTAAAGT
 GCTGTGGAGTAGTATATTTCACTGACTGGTTGGAATGACAGAGATGGACTGGCCCCCAGAT
 TCCTGCTGTGTTAGAGAATTCCCAGGATGTTCCAAACAGGCCCACCAGGAAGATCTCAGTGA
 CCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTTTTTTGAGAGGAACCAAACAACCTGC
 AGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAATCCTGGCCATGATTCTCACC
 ATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGGACAGACCAAATGATGTC
 CTTGAAGAATGACAACTCTCAGCACCTGTCAATGTCCCTCAGTAGAACTGTTGAAACCAAGCC
 TGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACTTTGAGATGGAG
 GAGTTA**TAAAA**AGAAATGTCACAGAAGAAAACCAAACTTGTTTTATTGGACTTGTGAATT
 TTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAATAACA
 CCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTCAC
 CACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGC
 CTGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTTC
 CGCATCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTT
 CTACCAACTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATA
 ACTTTTATTACTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTC
 AATATTGGTGACTACCTAAATGTGATTTTTTGCTGGTTACTAAAATATTCTTACCACTTAAAA
 GAGCAAGCTAACACATTGTCTTAAGCTGATCAGGGATTTTTTTGTATATAAGTCTGTGTTAAA
 TCTGTATAATTTCAGTCGATTTTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAA
 ATTTGTCTGTATAGCATCATTATTTTTAGCCTTTCCTGTTAATAAAGCTTTACTATTCTGT
 CCTGGGCTTATATTACACATATAACTGTTATTTAAATACTTAACCACTAATTTTGAAAATTA
 CCAGTGTGATACATAGGAATCATTATTCAGAAATGTAGTCTGGTCTTTAGGAAGTATTAATAA
 GAAAATTTGCACATAACTTAGTTGATTCAGAAAGGACTTGTATGCTGTTTTTCTCCCAAATG
 AAGACTCTTTTTGACACTAAACACTTTTTAAAAAGCTTATCTTTGCCTTCTCCAAACAAGAA
 GCAATAGTCTCCAAGTCAATATAAATCTACAGAAAATAGTGTTCTTTTTCTCCAGAAAAAT
 GCTTGTGAGAATCATTAACACATGTGACAATTTAGAGATTCTTTGTTTTATTTCACTGATTA
 ATATACTGTGGCAAATTACACAGATTATTAATTTTTTTACAAGAGTATAGTATATTTATTT
 GAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTGTTTATTCTCAGAATATGGAA
 AGAAAATTAAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 154

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68880
```

```
><subunit 1 of 1, 305 aa, 1 stop
```

><MW: 35383, pI: 5.99, NX(S/T): 0

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPV
VHPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQ
WSDMVTCLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREF
PGCSKQAHQEDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWAL
YYDRREPGLTDQMMSLKNDNSQHLSCPSVELLKPSSLRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 155

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCC
TGCACCTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGA
CCTGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAG
ACCGGGAGGATCACAGAGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGC
CTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGAT
CCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTTGTGGTTGTCTCATCAAGG
TGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAG
CTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTT
CCCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGG
ACTCGGCCACAGGGAACTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAG
ACAGCCTGTAGGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGA
TGTGTGTGAAATCACAGAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTC
TCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCTGTGGGAAGAGCCTGAAGACCCCCCGT
GTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGGTGAGCATCCAGTACGA
CAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACT
GCTTCAGGAAACATAACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGC
AGCTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAA
AGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGC
CCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGA
TGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCA
GGTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAAGA
TGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCC
CTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGG
GGGCCCAGACCCCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATG
TCTGGAAGGCTGAGCTGT**TAA**TGCTGCTGCCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTG
CCCTGCCCACCTGGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCC
CTCTGCCCACAGCCTCAGCATTTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCC
TCGCAGCCCAGAGGCGCCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCC
AGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAA
GGAACTTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTA AAAAGCCCAGATCACTGTGGG
CTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAAGCCTA
CTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTT
ACCTACTGTTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCT
CTGGCAAAAAAAAAAAAAA

FIGURE 156

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68885

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47644, pI: 5.18, NX(S/T): 2

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLIKVILDKYYF
LCGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWF
SACFDNFTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSL
HCLACGKSLKTPRVVGGEESVDSWPWQVSIQYDKQHVCSSILDPHWVLTAAHCFRKHTDV
FNWKVRAGSDKLGSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFD
EELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPE
GGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

Downloaded from www.oxfordjournals.org/

FIGURE 157

GGGCTGAGGCACTGAGAGACCGGAAAGCCTGGCATTCCAGAGGGAGGGAAACGCAGCGGCATCCCCAGGCTCCAG
 AGCTCCCTGGTGACAGTCTGTGGCTGAGCCATGGCCCTCCCAGCCCTGGGCCTGGACCCCTGGAGCCTCCTGGGCC
 TTTTCCTCTTCCAACCTGCTTCAGCTGCTGCTGCCGACGACGACCGCGGGGGAGGCGGGCAGGGGGCCATGCCCA
 GGGTCAGATACTATGCAGGGGATGAACGTAGGGCACTTAGCTTCTTCCACCAGAAGGGCCTCCAGGATTTTGACA
 CTCTGCTCCTGAGTGGTGATGGAATACTCTCTACGTGGGGGCTCGAGAAGCCATTCTGGCCTTGATATCCAGG
 ATCCAGGGGTCCCCAGGCTAAAGAACATGATACCGTGGCCAGCCAGTGACAGAAAAAGAGTGAATGTGCCTTTA
 AGAAGAAGAGCAATGAGACACAGTGTTCAACTTCATCCGTGTCTGGTTTTCTTACAATGTCACCCATCTCTACA
 CCTGCGGCACCTTCGCCTTCAGCCCTGCTTGTACCTTCATTGAACTTCAAGATTCTACCTGTTGCCATCTCGG
 AGGACAAGGTCATGGAGGGAAAAGGCCAAAGCCCCTTTGACCCCGCTCACAAGCATACGGCTGTCTTGGTGGATG
 GGATGCTCTATTCTGGTACTATGAACAACCTTCTGGGCAGTGAGCCCATCTGATGCGCACACTGGGATCCCAGC
 CTGTCTCAAGACCGACAACCTTCTCCGCTGGCTGCATCATGACGCCTCCTTTGTGGCAGCCATCCCTTCGACCC
 AGGTCGTCTACTTCTTCTTCGAGGAGACAGCCAGCGAGTTTGACTTCTTTGAGAGGCTCCACACATCGCGGGTGG
 CTAGAGTCTGCAAGAATGACGTGGGCGGCGAAAAGCTGCTGCAGAAGAAGTGGACCACCTTCCTGAAGGCCCAGC
 TGCTCTGCACCCAGCCGGGGCAGCTGCCCTTCAACGTATCCGCCACGCGGTCTGCTCCCCGCCGATTCTCCCA
 CAGCTCCCCACATCTACGAGTCTTCACCTCCAGTGGCAGGTTGGCGGGACCAGGAGCTCTGCGGTTTGTGCCT
 TCTCTCTCTTGGACATTGAACGTGTCTTTAAGGGGAAATACAAAGAGTTGAACAAAGAACTTCACGCTGGACTA
 CTTATAGGGGCCCTGAGACCAACCCCCGGCCAGGCAGTTGCTCAGTGGGCCCCCTCCTCTGATAAGGCCCTGACCT
 TCATGAAGGACCATTTCTGATGGATGAGCAAGTGGTGGGGACGCCCTGCTGGTGAAATCTGGCGTGGAGTATA
 CACGGCTTGCACTGGAGACAGCCAGGGCCTTGATGGGCACAGCCATCTTGTATGTACCTGGGAACCACCACAG
 GGTGCTCCACAAGGCTGTGGTAAGTGGGGACAGCAGTGCTCATCTGGTGGAAAGAGATTGAGCTGTTCCCTGACC
 CTGAACCTGTTGCAACCTGCAGCTGGCCCCCAGGAGTGCAGTGTGTTGTAGGCTTCTCAGGAGGTGTCTGGA
 GGGTGCCCCGAGCCAACCTGTAAGTGTCTATGAGAGCTGTGTGGACTGTGTCCTTGCCCCGGACCCCCACTGTGCCT
 GGGACCTGAGTCCCGAACCTGTTGCCTCCTGTCTGCCCCAACCTGAACTCCTGGAAGCAGGACATGGAGCGGG
 GGAACCCAGAGTGGGCATGTGCCAGTGGCCCCATGAGCAGGAGCCTTCGGCCTCAGAGCCGCCCGCAAATCATT
 AAGAAGTCTGGCTGTCCCCAACTCCATCCTGGAGCTCCCTGCCCCACCTGTCAGCCTTGGCCTCTTATTATT
 GGAGTCATGGCCCAGCAGCAGTCCCAGAAGCCTCTTCCACTGTCTACAATGGCTCCCTCTTGCTGATAGTGCAGG
 ATGGAGTTGGGGTCTCTACAGTGTGGGCAACTGAGAAATGGCTTTTCATACCCTGTGATCTCCTACTGGGTGG
 ACAGCCAGGACCAGACCCTGGCCCTGGATCCTGAACTGGCAGGCATCCCCCGGGAGCATGTGAAGGTCCCCTTGA
 CCAGGGTCAGTGGTGGGGCCGCCCTGGCTGCCAGCAGTCTACTGGCCCCACTTTGTCACTGTCACTGTCTCT
 TTGCCTTAGTGCTTTTCAGGAGCCCTCATCATCCTCGTGGCCTCCCCATTGAGAGCACTCCGGGCTCGGGGCAAGG
 TTCAGGGCTGTGAGACCCTGCGCCCTGGGGAGAAGCCCCGTTAAGCAGAGAGCAACACCTCCAGTCTCCAAGG
 AATGCAGGACCTCTGCCAGTGATGTGGACGCTGACAACAACCTGCCTAGGCACTGAGGTAGCTTAAACTCTAGGCA
 CAGGCCGGGGCTGCGGTGCAGGCACCTGGCCATGCTGGCTGGGCGGCCCAAGCACAGCCCTGACTAGGATGACAG
 CAGCACAAAAGACCACCTTTCTCCCTGAGAGGAGCTTCTGCTACTCTGCATCACTGATGACACTCAGCAGGGTG
 ATGCACAGCAGTCTGCCTCCCTATGGGACTCCCTTCTACCAAGCACATGAGCTCTCTAACAGGGTGGGGGCTAC
 CCCCAGACCTGCTCCTACACTGATATTGAAGAACCTGGAGAGGATCCTTCAGTTCTGGCCATTCCAGGGACCCTC
 CAGAAACACAGTGTTCAGAGACCCTAAAAAACCTGCCTGTCCCAGGACCCTATGGTAATGAACACCAAACATC
 TAAACAATCATATGCTAACATGCCACTCCTGGAACTCCACTCTGAAGCTGCCGCTTTGGACACCAACACTCCCT
 TCTCCCAGGGTCATGCAGGGATCTGCTCCCTCCTGCTTCCCTTACCAGTCGTGCACCGCTGACTCCCAGGAAGTC
 TTTCTGAAGTCTGACCACCTTTCTTCTTGCTTCAGTTGGGGCAGACTCTGATCCCTTCTGCCCTGGCAGAATGG
 CAGGGGTAATCTGAGCCTTCTTCACTCCTTTACCCTAGCTGACCCCTTCACCTCTCCCCCTCCCTTTTCTTTGT
 TTTGGGATTGAGAAAACCTGCTTGTGAGAGACTGTTTATTTTTATTAAAAATATAAGGCTTAAAAAA

FIGURE 158

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71166
><subunit 1 of 1, 761 aa, 1 stop
><MW: 83574, pI: 6.78, NX(S/T): 4
MALPALGLDPWSLLGLFLFQLLQLLLPTTTAGGGGQGMPRVRYAGDERRALSFFHQKGLQ
DFDTLLLSGDGNTLYVGAREAILALDIQDPGVPR LKNMIPWPASDRKKSECAFKKKSNETQC
FNFIRVLVSYNVTHLYTCGTFAFSPACTFIELQDSYLLPISEDKVMEGKGQSPFDPAHKHTA
VLVDGMLYSGTMNNFLGSEPILMRTLGSQPVLKTDNFLRWLHHDASFVAAIPSTQVVYFFFE
ETASEFDDFFERLHTSRVARVCKNDVGGEKLLQKKWTTFLKAQLLCTQPGQLPFNVIRHAVLL
PADSPTAPHIYAVFTSQWQVGGTRSSAVCAFSLLDIERVFKGKYKELNKETSRWTTYRGPET
NPRPGSCSVGPSSDKALTFMKDHFLMDEQVVGTPLLVKSGVEYTRLAVETAQGLDGHSHLVM
YLGTTTGS LHKAVVSGDSSAHLVEEIQLFPDPEPVRNLQLAPTQGAVFVGFSGGVWRVPRAN
CSVYESCVDCVLARDPHCAWDPESTRCCLLSAPNLNSWKQDMERGNPEWACASGPMSRSLRP
QSRPQIIKEVLAVPNSILELPCPHLSALASYWSHGPAAVPEASSTVYNGSLLLIVQDGVGG
LYQCWATENGFSYPVISYWVDSQDQTLALDPELAGIPREHVKVPLTRVSGGAALAAQQSYWP
HFVTVTVL FALVLSGALIILVASPLRALRARGKVQGCETLRPGEKAPLSREQHLQSPKECRT
SASDVDADNNCLGTEVA

Signal peptide:

amino acids 1-30

Transmembrane domains:

amino acids 136-156, 222-247, 474-490, 685-704

FIGURE 159

AGGGTCCCTTAGCCGGGCGCAGGGCGCGCAGCCCAGGCTGAGATCCGCGGCTTCCGTTAGAAG
 TGAGC**ATG**GCTGGGCAGCGAGTGCTTCTTCTAGTGGGCTTCCTTCTCCCTGGGGTCTTGCTC
 TCAGAGGCTGCCAAAATCCTGACAATATCTACAGTAGGTGGAAGCCATTATCTACTGATGGA
 CCGGGTTTCTCAGATTCTTCAAGATCACGGTCATAATGTCACCATGCTTAACCACAAAAGAG
 GTCCTTTTATGCCAGATTTTAAAAAGGAAGAAAAATCATATCAAGTTATCAGTTGGCTTGCA
 CCTGAAGATCATCAAAGAGAATTTAAAAAGAGTTTGTATTTCTTTCTGGAAGAACTTTAGG
 TGGCAGAGGAAAATTTGAAAACCTATTAAATGTTCTAGAATACTTGGCGTTGCAGTGCAGTC
 ATTTTTTAAATAGAAAGGATATCATGGATTCCCTTAAAGAATGAGAACTTCGACATGGTGATA
 GTTGAAACTTTTGACTACTGTCCTTTCCTGATTGCTGAGAAGCTTGGGAAGCCATTGTGGC
 CATTCTTCCACTTCATTTCGGCTCTTTGGAATTTGGGCTACCAATCCCCTTGTCTTATGTTT
 CAGTATTCGGTTCCCTTGCTGACTGATCACATGGACTTCTGGGGCCGAGTGAAGAATTTTCTG
 ATGTTCTTTAGTTTCTGCAGGAGGCAACAGCACATGCAGTCTACATTTGACAACACCATCAA
 GGAACATTTACAGAAGGCTCTAGGCCAGTTTGTCTCATCTTCTACTGAAAGCAGAGTTGT
 GGTTCATTAACTCTGACTTTGCCTTTGATTTTGTCTCGACCTCTGCTTCCCAACACTGTTTAT
 GTTGGAGGCTTGATGGAAAAACCTATTAAACCAGTACCACAAGACTTGGAGAACCTTCATTGC
 CAAGTTTGGGGACTCTGGTTTTGTCTTGTGACCTTGGGCTCCATGGTGAACACCTGTCAGA
 ATCCGGAAATCTTCAAGGAGATGAACAATGCCTTTGCTCACCTACCCCAAGGGGTGATATGG
 AAGTGTCAAGTGTCTCATTGGCCCCAAGATGTCCACCTGGCTGCAAATGTGAAAATTGTGGA
 CTGGCTTCCTCAGAGTGACCTCCTGGCTCACCCAAGCATCCGTCTGTTTGTCACCCACGGCG
 GGCAGAATAGCATAATGGAGGCCATCCAGCATGGTGTGCCCATGGTGGGGATCCCTCTCTTT
 GGAGACCAGCTGAAAACATGGTCCGAGTAGAAGCCAAAAAGTTTGGTGTCTTCTATTAGTT
 AAAGAAGCTCAAGGCAGAGACATTGGCTCTTAAGATGAAACAAATCATGGAAGACAAGAGAT
 ACAAGTCCGCGGCAGTGGCTGCCAGTGTATCCTGCGCTCCACCCGCTCAGCCCCACACAG
 CGGCTGGTGGGCTGGATTGACCACGTCTCCAGACAGGGGGCGCGACGCACCTCAAGCCCTA
 TGTCTTTCAGCAGCCCTGGCATGAGCAGTACCTGTTTCGACGTTTTTGTGTTTCTGCTGGGGC
 TCACTCTGGGGACTCTATGGCTTTGTGGGAAGCTGCTGGGCATGGCTGTCTGGTGGCTGCGT
 GGGGCCAGAAAGGTGAAGGAGACAT**TAA**GGCCAGGTGCAGCCTTGGCGGGGTCTGTTTGGTGG
 GCGATGTCACCATTTCTAGGGAGCTTCCCACTAGTTCTGGCAGCCCCATTCTCTAGTCCTTC
 TAGTTATCTCCTGTTTTCTTGAAGAACAGGAAAAATGGCCAAAAATCATCCTTCCACTTGC
 TAATTTTGCTACAAATTCATCCTTACTAGCTCCTGCCTGCTAGCAGAAATCTTCCAGTCTT
 CTTGTCTCCTTTGTTTGCCATCAGCAAGGGCTATGCTGTGATTCTGTCTCTGAGTGACTTG
 GACCACTGACCCTCAGATTTCCAGCCTTAAAAATCCACCTTCCTTCTCATGCGCCTCTCCGAA
 TCACACCCTGACTCTTCCAGCCTCCATGTCCAGACCTAGTCAGCCTCTCTCACTCCTGCCCC
 TACTATCTATCATGGAATAACATCCAAGAAAGACACCTTGCATATTCTTTAGTTTCTGTTT
 TGTCTCCACATATTCTCTTCAATGCTCAGGAAGCCTGCCCTGTGCTTGAGAGTTCAGGGC
 CGGACACAGGCTCACAGGTCTCCACATTGGGTCCCTGTCTCTGGTGCCACAGTGAGCTCCT
 TCTTGGCTGAGCAGGCATGGAGACTGTAGTTTCCAGATTTCCCTGAAAAATAAAAGTTTACA
 GCGTTATCTCTCCCAACCTCACTAA

FIGURE 160

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71169

><subunit 1 of 1, 523 aa, 1 stop

><MW: 59581, pI: 8.68, NX(S/T): 1

MAGQRVLLLVGFLLPGVLLSEAAKILTISTVGGSHYLLMDRVSQILQDHGHNVTMLNHKRGPF
FMPDFKKEEKSQVISWLAPEDHQREFKKSFDFFLEETLGGRGKFENLLNVLEYLALQCSHF
LNRKDIMDSLKNENFDMVIVETFDYCPFLIAEKLKGPFVAILSTSFGSLEFGLPIPLSYVPV
FRSLTLDHMDFWGRVKNFLMFFSFCRRQQHMQSTFDNTIKEHFTEGSRPVLSHLLLKAELWF
INSDFAFDFARPLLNTVYVGGLMEKPIKVPQDLENFIAKFGDSGFVLVTLGSMVNTCQNP
EIFKEMNNAFAHLPQGVWKCQCASHWPKDVHLAANVKIVDWLPQSDLLAHPSIRLFVTHGGQ
NSIMEAIOHGVPMVGIPFLFGDQPENMVRVEAKKFGVSIQLKKLKAETLALKMKQIMEDKRYK
SAAVAASVILRSHPLSPTQRLVGWIDHVLQTGGATHLKPYVFQQPWHEQYLFDFVFLGLT
LGTLLWLCGKLLGMAVWWLRGARKVKET

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 483-504

FIGURE 161

GGGCTGTTGATTGTGGGGGATTTTGAAGAGAGGAGGAATAGGAGGAAGGGGTTGAGGGGCT
GCCTCTGGCATATGCACACACTCACACATTCTGTACACCCGTCACACACACATACCATGTT
CTCCATCCCCCAGGTCCAGCCCTCAGTGCTGTCCCATCCAGCAGGGGCTACCCTGAAGCTCT
GGCTGCAGCCCTCCCGTCCAGTGGGCAGGCGGCTTCATCCCTCCTTTCTCTCCCAAAGCCCA
ACTGCTGTCACTGCATGCTCTGCCAAGGAGGAGGGAACTGCAGTGACAGCAGGAGTAAGAGT
GGGAGGCAGGACAGAGCTGGGACACAGGTATGGAGAGGGGGTTTCAGCGAGCCTAGAGAGGGC
AGACTATCAGGGTGCCGGCGGTGAGAATCCAGGGAGAGGAGCGGAAACAGAAGAGGGGGCAGT
AGACCGGGGCACTTGTGGGTTCAGAGCCCCCTCAGCCATGTTGGGAGCCAAGCCACACTGGC
TACCAGGTCCCCTACACAGTCCCGGGCTGCCCTTGGTTCTGCTGCTTCTGGCCCTGGGGGCC
GGGTGGGCCCAGGAGGGGTTCAGAGCCCGTCTGCTGGAGGGGGAGTGCCTGGTGGTCTGTGA
GCCTGGCCGAGCTGCTGCAGGGGGGGCCCGGGGAGCAGCCCTGGGAGAGGCACCCCTGGGC
GAGTGGCATTGTCTGCGGTCCGAAGCCACCACCATGAGCCAGCAGGGGAAACCGGCAATGGC
ACCAGTGGGGCCATCTACTTCGACCAGGTCTGCTGTAACGAGGGCGGTGGCTTTGACCGGGC
CTCTGGCTCCTTCGTAGCCCCCTGTCCGGGGTGTCTACAGCTTCCGGTTCCATGTGGTGAAGG
TGTACAACCGCCAAACTGTCCAGGTGAGCCTGATGCTGAACACGTGGCCTGTCTCTCAGCC
TTTGCCAATGATCCTGACGTGACCCGGGAGGCAGCCACCAGCTCTGTGCTACTGCCCTTGGA
CCCTGGGGACCGAGTGTCTCTGCGCCTGCGTCGGGGGAATCTACTGGGTGGTTGGAAATACT
CAAGTTTCTCTGGCTTCCTCATCTTCCCTCTCTGAGGACCCAAGTCTTTCAAGCACAGAAT
CCAGCCCCCTGACAACTTTCTTCTGCCCTCTCTTGCCCCAGAAACAGCAGGGCAGGAGAGAG
ACTCCCTCTGGCTCCTATCCCACCTCTTTGCATGGGACCCTGTGCCAAACACCCAAGTTTAA
GAGAAGAGTAGAGCTGTGGCATCTCCAGACCAGGCCTTTCCACCCACCCACCCCAAGTTTAA
CTCCCAGCCACCTGCTGCATCTGTTCTGCTGAGCCCTAGGATCAGGGCAAGGTTTGGCA
AGAAGGAAGATCTGCACTACTTTGCGGCCTCTGCTCCTCCGGTTCCTCCACCCAGCTTCCT
GCTCAATGCTGATCAGGGACAGGTGGCGCAGGTGAGCCTGACAGGCCCCACAGGAGCCCA
ATGGACAAGCCTCAGCGTACCCTGCAGGCCTTCTTCTGAGGAAAGCCAGCATCAGCGATC
TCAGCCAGCACCGTCAGAAGCTGAGCCAGCACCGTATGGGCTAGGGTGGGAGGCTCAGCCAC
AGGCAGAAGGGTGGGAAGGGCCTGGAGTCTGTGGCTGGTGAGGAAGGAAGGAGGGTGTATTG
TCTAGACTGAACATGGTACACATTCTGCATGTATAGCAGAGCAGCCAGCAGGTAGCAATCCT
GGCTGTCTTCTATGCTGGATCCCAGATGGACTCTGGCCCTTACCTCCCCACCTGAGATTAG
GGTGAGTGTGTTTGTCTCTGGCTGAGAGCAGAGCTGAGAGCAGGTATACAGAGCTGGAAGTGG
ACCATGGAAAACATCGATAACCATGCATCCTCTTGCTTGGCCACCTCCTGAAACTGCTCCAC
CTTTGAAGTTTGAACTTTAGTCCCTCCACACTCTGACTGCTGCCTCCTTCTCCAGCTCTC
TCACTGAGTTATCTTCACTGTACCTGTTCCAGCATATCCCCACTATCTCTCTTTCTCCTGAT
CTGTGCTGTCTTATTCTCCTCCTTAGGCTTCCTATTACCTGGGATTCCATGATTCATTCTT
CAGACCCTCTCCTGCCAGTATGCTAAACCCTCCCTCTCTCTTTCTTATCCCGCTGTCCCATT
GGCCCAGCCTGGATGAATCTATCAATAAAACAACCTAGAGAATGGTGGTCAGTGAGACACTAT
AGAATTACTAAGGAGAAGATGCCTCTGGAGTTTGGATCGGGTGTACAGGTACAAGTAGGTA
TGTTGCAGAGGAAAATAAATATCAAACCTGTATACTAAAATTAAAAA

FIGURE 162

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71180

><subunit 1 of 1, 205 aa, 1 stop

><MW: 21521, pI: 7.07, NX(S/T): 1

MLGAKPHWLPGPLHSPGLPLVLVLLALGAGWAQEGSEPVLLERGECLVVCEPGRAAAGGPGGA
ALGEAPPGRVAFAAAVRSHHHEPAGETGNGTSGAIYFDQVLVNEGGGFDRASGSFVAPVRGVY
SFRFHVVVKVYNRQTVQVSLMLNTWPVISAFANDPDVTREAATSSVLLPLDPGDRVSLRLRRG
NLLGGWKYSSFSGFLLIFPL

Signal peptide:

amino acids 1-32

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 163

[illegible]

FIGURE 164

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71184
><subunit 1 of 1, 388 aa, 1 stop
><MW: 43831, pI: 9.64, NX(S/T): 3
MKTLLIAAYSGVLRGERQAEADRSQRSHGGPALSREGSGRWGTGSSILSALQDLFSVTWLNRS
KVEKQLQVISVLQWVLSFLVLGVACSAILMYIFCTDCWLI AVL YFTWL VFDWNTPKKGRRS
QWVRNWAVWRYFRDYFPIQLVKTHNLLTTRNYIFGYHPHGIMGLGAF CNFSTEATEVSKKFP
GIRPYLATLAGNFRMPVLREYLMSGGICPVSRDTIDYLLSKNGSGNAIIIVVGAAESLSSM
PGKNAVTLRNRKGFVKLALRHGADLVPIYSFGENEVYQVIFEEGSWGRWVQKKFQKYIGFA
PCIFHGRGLEFSSDTWGLVPYSKPITTVVGEPITIPKLEHPTQQDIDLYHTMYMEALVKLFDK
HKTKFGLPETEVLEVN

Important features of the protein:**Transmembrane domain:**

amino acids 76-97

N-glycosylation sites.

amino acids 60-63, 173-176, 228-231

N-myristoylation sites.

amino acids 10-15, 41-46, 84-89, 120-125, 169-174, 229-234, 240-
245, 318-323, 378-383

FIGURE 165

[illegible]

FIGURE 166

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71213
><subunit 1 of 1, 368 aa, 1 stop
><MW: 42550, pI: 9.11, NX(S/T): 1
MGLLAFLKTQFVLHLLVGFVVFVVSGLVINQVQLCTLALWPVSKQLYRRLNCRSLAYSLWSQLV
MLLEWWSCTECTLFQATVERFGKEHAVIILNHNFEIDFLCGWTMCERFGVLGSSKVLAKK
ELLYVPLIGWTWYFLEIVFCKRKWEEDRDTVVEGLRRLSDYPEYMWFLLYCEGTRFTETKHR
VSMEVAAAKGLPVLKYHLLPRTKGFTTAVKCLRGTVAAVYDVTLLNFRGNKNPSLLGILYGKK
YEADMCVRRFPLEDIPLDEKEAAQWLHKLYQEKDALQEIYNQKGMFPGEQFKPARRPWTLLN
FLSWATILLSPLFSFVLGVFASGSPLLILTFGLGFVGAASFGVRRLIGESLEPGRWRLQ
```

Important features of the protein:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 307-323, 335-352

Tyrosine kinase phosphorylation sites.

amino acids 160-168, 161-169

FIGURE 167

GATATTCTTTATTTTAAAGAATCTGAAGTACTATGCATCACTCCCTCCAATGTCCTGGGGCA
GCCACCAGGCATATTCATCTTTGTGTGTGTTTTCTTTGCTTTAGCACTGGGGCACTTCTT
GCTTATTTCTTTGGTAGGAAAGGGGCTCAGTTTGTCTTGTGGGGTTGGTGGCAGGCAGGCCG
GCTTACGCCTGATACGGCCCTGGGTAGAAAGGGAAGGGAAGATAAACTTTTATACAAATGGG
GATAGCTGGGGTCTGAGACCTGCTTCCTCAGTAAAATTCCTGGGATCTGCCTATACCTTCTT
TTCTCTAACCTGGCATAACCCTGCTTAAAGCCTCTCAGGGCTTCTCTCTGTTCTTAGGATCAA
AGTATTTAGAGCTACAAGAGCCCTCATGGTCTGGCCCCTGCCCCCTGGCCAGCTTCATTGT
ACATGTGGTGTTCTCTTGTCGTTCCCTGTAATGTGGTATGCCATGGGGTCTTTGCACAAGCCT
TTCCTCTTTGGCTGGACACTGTTCCCTGCCCCCCCCATACTCTTCCTACTTAATATGTAGTC
ATCCTGCAGATTTCAATTCTAACATCATTTTCTCCAGGGATCCTGGCCTGACAGAATCTCAT
CTTGTTTAAATGCTCTCATAAGACCACTTGTTTCCCTTTTGCAGCACTTGCCACTCAGTTGTA
TCTTTATGTGCGTTTGTGGTTGTATGGGTTGTGTCTGTTCCCCAGAATGCCAGCTCTGAGC
TGCGTGAGGGTCAAGGGCATTGCTGTGCCTGCCAGGTATAGTGCCTACATGTGGTGGGTGCT
CATGTTTTAGAGACTAAATGGAGGAGGAGATGAGGAAAAGATTGAAATCTCTCAGTTCACCA
GATGGTGTAGGGCCCAGCATTGTAAATTCACACGTTGACTGTGCTTGTGAATTATCTGGGGA
TGCAGGTCCTGATTCAGTAGGCCAGGTTGGGCATCTCTAACAACTCCCACGTGATGCTGA
TGCTGGTCCTATGAACTATACTAAATAGTAAGAATCTATGGAGCCAGGCTGGGCATGGTGGC
TCACACCTATGATCCCAGCACTTTGGGAGGCTGAGGCAGGCTGATCACCTGGAGTCAGGATT
TCAAGACTAGCCTGGCCAACATGGTGGAAACCCCATCTGTACTAAAAATACACAAATTAGCTG
GGCATGGTGGCACATGCCTGTAGTCCCAGCTACTTGGGAGGCTGAAGCAAGAGAATCGCTTG
AACCTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCAGGCCACTGTATTCCAACCAGGGTGAC
AGAGTGAGACTCTATGTCCAAAAAAAAAAAAA

FIGURE 168

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71234
><subunit 1 of 1, 143 aa, 1 stop
><MW: 15624, pI: 9.58, NX(S/T): 0
MHHSLQCPGAATRHHLCVCFSSFALALGHFLLISLVGKGLSLSCGVGGRQAGLRLIRPWVRR
EGKINFYTNNGDSWGLRPASSVKFLGSAYTFFSLTWHTLLKASQGFSLFLGSKYLELQEPSWS
GPCPPGQLHCTCGVLLSFL
```

Important features of the protein:

Signal peptide:

amino acids 1-28

1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200

FIGURE 169

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGA
 TTATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAGCCAC
 AGGAGTTGAACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTAC
 ATTAAAATCTGTTTTTTTGTCTCTTGTAAGTAGCCTTTACCTTCCTAACACAGAGGATCTGT
 CACTGTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCAC
 ACCGTCCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTC
 ACCAACTGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGA
 GCTGCGGCTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTT
 CATAGGCGATGGCTCCCACTGCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGG
 CCAGGACGGGCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCCGCCAT
 CTAACCTTTTCATGTCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGG
 AACCCAGAGCTTGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTT
 TATGTGACAGGACTTGCATTCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCA
 GGAAGGAACTTGTGCCAAATTATGGGTCAGAAAAGATGGAGGTGTTGGGTATCACAAGGC
 ATCGAGTCTCCTGCATTCAGTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACT
 CGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGATCCACGTACCAGCTGCTG
 AAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAG
 CCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCTCCTTCCCTC
 TGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATG
 GCTCAGTGTTGGCCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCG
 AACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCT
 CAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGT
 TCAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAA
 TCCAAACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATG
 CCAACATTTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTG
 GCCAACATGGTGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGT
 GTGTGCCTGTATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGT
 GAAGGAGGCTGAGACAGGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGA
 AAAAATAAAAAAAGAATTATGGTTATTTGTAA

FIGURE 170

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71277

><subunit 1 of 1, 109 aa, 1 stop

><MW: 11822, pI: 8.63, NX(S/T): 0

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTA
SPCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

172/249

FIGURE 171

GCGGGCCCGCGAGTCCGAGACCTGTCCCAGGAGCTCCAGCTCACGTGACCTGTCACTGCCTC
CCGCCGCCTCCTGCCCCGCGCC**ATG**ACCCAGCCGGTGCCCCGGCTCTCCGTGCCCCGCCGCGCT
GGCCCTGGGCTCAGCCGCACTGGGCGCCGCCTTCGCCACTGGCCTCTTCCTGGGGAGGCGGT
GCCCCCATGGCGAGGCGGCGAGAGCAGTGCCTGCTTCCCCCGAGGACAGCCGCCTGTGG
CAGTATCTTCTGAGCCGCTCCATGCGGGAGCACCCGGCGCTGCGAAGCCTGAGGCTGCTGAC
CCTGGAGCAGCCGAGGGGGATTCTATGATGACCTGCGAGCAGGCCAGCTCTTGGCCAACC
TGGCGCGGCTCATCCAGGCCAAGAAGGCGCTGGACCTGGGCACCTTCACGGGCTACTCCGCC
CTGGCCCTGGCCCTGGCGCTGCCCCGCGGACGGGCGCGTGGTGACCTGCGAGGTGGACGCGCA
GCCCCCGGAGCTGGGACGGCCCCTGTGGAGGCAGGCCGAGGCGGAGCACAAGATCGACCTCC
GGCTGAAGCCCGCCTTGGAGACCCTGGACGAGCTGCTGGCGGCGGGCGAGGCCGGCACCTTC
GACGTGGCCGTGGTGGATGCGGACAAGGAGAACTGCTCCGCCTACTACGAGCGCTGCCTGCA
GCTGCTGCGACCCGGAGGCATCCTCGCCGTCCTCAGAGTCCTGTGGCGCGGGAAGGTGCTGC
AACCTCCGAAAGGGGACGTGGCGGCCGAGTGTGTGCGAAACCTAAACGAACGCATCCGGCGG
GACGTCAGGGTCTACATCAGCCTCCTGCCCCCTGGGCGATGGACTCACCTTGGCCTTCAAGAT
CTAGGGCTGGCCCCCTAGTGAGTGGGCTCGAGGGAGGGTTGCCTGGGAACCCAGGAATTGAC
CCTGAGTTTTTAAATTCGAAAATAAAGTGGGGCTGGGACACAAAAAAAAAAAAAAAAAAAAA

FIGURE 172

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71282
><subunit 1 of 1, 262 aa, 1 stop
><MW: 28809, pI: 8.80, NX(S/T): 1
MTQPVPRLSVPAALALGSAALGAATGLFLGRRCPPWRGRREQCLLPEDSRLWQYLLSRS
MREHPALRSRLRLTLEQPQGDSMMTCEQAQLLANLARLIQAKKALDLGTFTGYSALALALAL
PADGRVVTCEVDAQPPPELGRPLWRQAEAEHKIDLRRLKPALETLDELLAAGEAGTFDVAVVDA
DKENCSAYYERCLQLLRPGGILAVLRVLWRGKVLQPPKGDVAAECVRNLNERIRRDVRVYIS
LLPLGDGLTLAFKI
```

Important features of the protein:**Signal peptide:**

amino acids 1-25

Transmembrane domains:

amino acids 8-30, 109-130

N-glycosylation site.

amino acids 190-193

Tyrosine kinase phosphorylation site.

amino acids 238-246

N-myristoylation sites.

amino acids 22-27, 28-33, 110-115, 205-210, 255-260

Amidation sites.

amino acids 31-34, 39-42

FIGURE 173

[illegible]

FIGURE 174

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71286

><subunit 1 of 1, 671 aa, 1 stop

><MW: 74317, pI: 7.61, NX(S/T): 0

MPHAFKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPKNKYPIFFFFGTHETAFLGPKDLFPYD
KCKDKYGKPNKRKGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEANPADGSDADEDDERD
VMAVTAVTATAASDRMESDSDSDKSSDNSGLKRKTPALKMSVSKRARKASSDLDQASVSPSE
EENSESSSESEKTSQDFTPEKKA AVRAPRRGPLGGRKKKKAPSASDSDSKADSDGAKPEPV
AMARSASSSSSSSSSSSDSDSVKPKPPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDEV
DRISEWKRRDEARRRELEARRRREQEEELRRLREQEKEEKERRRERADRGEAERGSGGSSGD
ELREDDEPVKKRGRKGRGRGPPSSSDSEPEAELEAREAKKSAKKPQSSSTEPARKPGQKEKRV
RPEEKQQAQPVKVERTKRSEGFSDMRKVEKKKEPSVEEKLQKLHSEIKFALKVDS PDVKRC
LNALEELGTLQVTSQILQKNTDVVATLKKIRRYKANKDVMEKAAEVYTRLKSRVLGPKIEAV
QKVNKAGMEKEKAEKLAGEEELAGEEAPQEKAEDKPSTDLSAPVNGEATSQKGESAEDKEHE
EGRDSEEGPRCGSSEDLHDSVREGPDLD RPGSDRQERERARGDSEALDEES

Signal peptide:

amino acids 1-13

GTTGGTTCTCCTGGATCTTTCACCTTACCAACTGCAGACCTTGGGACTCATCAGCCTCAATAATTATATTAATTAATTA
 ACACCATTTGAAAGAGAACATTGTTTTTCATCATGAATGCTAATAAAGATGAAAGACTTAAAGCCAGAAGCCAAGA
 TTTTCACCTTTTTCTGCTTTGATGATGCTAAGCATGACCATGTTGTTTTCTTCCAGTCACTGGCACTTTGAAGCA
 AAATATTCCAAGACTCAAGCTAACCTACAAAGACTTGTCTGTTTTCAAATAGCTGTATTCCCTTTTTGGGTTTCATC
 AGAAGGACTGGATTTTTCAAACCTTTCTCTTAGATGAGGAAAGAGGCAAGGCTGCTCTTGGGAGCCAAAGACCACAT
 CTTTCTACTCAGTCTGGTTGACTTAAACAAAAATTTTAAAGAAATTTATTGGCCTGCTGCAAGGAACGGGTGGA
 ATTATGTAAATTAGCTGGGAAAGATGCCAATACAGAATGTGCAAATTTTCATCAGAGTACTTCAGCCCTATAACAA
 AACTCACATATATGTGTGTGGAAGTGGAGCATTTCATCCAATATGTGGGTATATTGATCTTGGAGTCTACAAGGA
 GGATATTATATTCAAACCTAGACACACATAATTTGGAGTCTGGCAGACTGAAATGTCCTTTTCGATCCTCAGCAGCC
 TTTTGTCTCAGTAATGACAGATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAAGATACTGCATT
 CACTCGATCCCTTTGGGCTACTCATGACCACCACATACAGAACTGACATTTTCAGAGCACTACTGGCTCAATGG
 AGCAAAATTTTATTGGAACCTTTCTTCATACCAGACACCTACAATCCAGATGATGATAAAATATTATTTCTTCTTCG
 TGAATCATCTCAAGAAGGCAGTACCTCCGATAAAACCATCCTTTCTCGAGTTGGAAGAGTTTGTAAAGATGATGT
 AGGAGGACAACGCAGCCTGATAAAACAAGTGGACGACTTTTCTTAAAGGCCAGACTGATTTGCTCAATTCCTGGAAG
 TGATGGGGCAGATACTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCCACAAGAGATGAAAGAAATCCTGT
 AGTATATGGAGTCTTTACTACAACCAGCTCCATCTTCAAAGGCTCTGCTGTTTGTGTGTATAGCATGGCTGACAT
 CAGAGCAGTTTTTAATGGTCCATATGCTCATAAGGAAAGTGACACCATCGTTGGGTGCAGTATGATGGGAGAAT
 TCCTTATCCACGGCCTGGTACATGTCCAAGCAAAACCTTATGCCACTGATTAAGTCCACCAGATATTTCCAGA
 TGATGTCATCAGTTTCATAAAGCGGCACCTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGAGGACCAACGTT
 CAAGAGAATCAATGTGGATTACAGACTGACACAGATAGTGGTGGATCATGTCAATGCAGAAGATGGCCAGTACGA
 TGTAATGTTTTCTTGGAAACAGACATTGGAACCTGTCCTCAAAGTTGTGAGCATTTCAAAGGAAAAGTGAATATGGA
 AGAGGTAGTGTCTGGAGGAGTTGCAGATATCCAAGCACTCATCAATCATCTTGAACATGGAATTTGTCTCTGAAGCA
 GCAACAATTGTACATTGGTTCCCGAGATGGATTAGTTCAGCTCTCTTGCACAGATCGCACACTTATGGGAAAGC
 TTGCGCAGACTGTTGTCTTGCCAGAGACCCCTACTGTGCTGGGATGGAAATGCATGCTCGATGATGCTCTTAC
 TTCTAAAAGGAGAGCTAGACGCCAAGATGTAAATATGGCGACCCAATCACCCAGTGTCTGGGACATCGAAGACAG
 CATTAGTCATGAAACTGCTGATGAAAAGGTGATTTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATACC
 TAAATCCCAACAGCAACTATTAAATGGTATATCCAGAGGTGAGGGGATGAGCATCGAGAGGAGTTGAAGCCCGA
 TGAAGAATCATCAAACCGAATATGGGCTACTGATTGCAAGTTTGCAGAAGAAGGATTCTGGGATGTATTACTG
 CAAAGCCCATGAGGACACACTTCTATCCACACCATGATGAAGCTGACTTTGAATGCTCATTGAGAATGAACAGATGGA
 AAATACCCAGAGGGCAGAGCATGAGGAGGGGCAGGTCAAGGATCTATTGGCTGAGTCACGGTTGAGATACAAAGA
 CTACATCCAAATCCTTAGCAGCCCAAACCTTCAGCCTCGACCAGTACTGCGAACAGATGTGGCACAGGGAGAAGCG
 GAGACAGAGAAACAAGGGGGGCCCAAAGTGGAAGCACATGCAGGAAATGAAGAAGAAACGAAATCGAAGACATCA
 CAGAGACCTGGATGAGCTCCCTAGAGCTGTAGCCACGTTAGTTTTCTACTTAATTTAAAGAAAAGAATTCCTTACC
 TATAAAAACATTGCTTCTGTTTTGTATATCCCTTATAGTAATTCATAAATGCTTCCCATTGGAGTTTTGCTAAGG
 CACAAGACAATAATCTGAATAAGACAATATGTGATGAATATAAGAAAGGCCAAAAAATTCATTGGAACGATTTT
 CCAAGAACAAATCTTGCACAAGCAAAGTATAAGAATTATCCTAAAAAATAGGGGGTTACAGTTGTAATGTTTTTA
 TGTTTTGAGTTTTGGAATTTATTGTGATGTAATAAGTTGAGCTAAGCAAGCCCCGAATTTGATAGTGTATAAGGT
 GCTTTATTCCCTCGAATGTCCATTAAAGCATGGAATTTACCATGCAGTTGTGCTATGTTCTTATGAACAGATATAT
 CATTCTTATTGGAACCCAGTACCTTTGTGGTAGGGAATAAGAGGTGAGACACAAATTAAGACAACCTCCCATTATC
 AACAGGAACTTTTCTCAGTGAGCCATTGCTCCTGGAGAATGGTATAGGAATTTGGAGAGGTGCATTATTTCTTCT
 TGGCCACTGGGGTTAAATTTAGTGTACTACAACATTGATTTACTGAAGGCCACTAATGTTTCCCCAGGATTTCT
 ATTGACTAGTCAGGAGTAACAGGTTTACAGAGAGAAGTTGGTGCTTAGTTATGTGTTTTTTAGAGTATATACTAA
 GCTCTACAGGACAGAAATGCTTAATAAATACTTTAATAAGATATGGGAAAATATTTTAATAAAACAAGGAAAACA
 TAATGATGTATAAATGCATCCTGATGGGAAGGCATGCAGATGGGATTTGTTAGAAGACAGAAGGAAAGACAGCCAT
 AAATCTGGCTTTTGGGAAACCTCATATCCCCATGAAAGGAAGAACAATCACAAATAAAGTGAGAGTAATGTAA
 TGGAGCTCTTTTCACTAGGTTATAAGTAGCTGCCAATTTGTAATTCATCTGTTAAAAAAAATCTAGATTATAACA
 AACTGCTAGCAAAATCTGAGGAAAACATAAATCTTCTGAAGAATCATAGGAAGAGTAGACATTTTATTTATAACC
 AATGATATTTTCAGTATATATTTTCTCTCTTTTAAAAAATATTTATCATACTCTGTATATTATTTCTTTTTACTGC
 CTTTATTCTCTCCTGTATATTGGATTTTGTGATTATATTTGAGTGAATAGGAGAAAACAATATATAACACACAGA
 GAATTAAGAAAATGACATTTCTGGGGAGTGGGGATATATATTTGTTGAATAACAGAACGAGTGTAAAATTTTAAC
 ACGGAAAGGGGTAAATTAACCTTTTGACATCTTCACTCAACCTTTTCTCATTGCTGAGTTAATCTGTTGTAATT
 GTAGTATTGTTTTTGTAAATTTAAACAATAAATAAGCCTGCTACATGT

FIGURE 176

```

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71883
><subunit 1 of 1, 777 aa, 1 stop
><MW: 89651, pI: 7.97, NX(S/T): 3
MNANKDERLKARSQDFHLFPALMMLSMTMLFLPVTGTLKQNIPLRLKLTyKDLLLSNSCIPFL
GSSEGLDFQTLILLDEERGRLLLGAkDHIFLLSLVDLNKNFKKIYWPAAKERVELCKLAGKDA
NTECANFIRVLQPYNKTHIYVCGTGAFHPICGYIDLGVYKEDIIFKLDTHNLESGRLKCPFD
PQQPFASVMTDEYLYSGTASDFLGKDTAFTRSLGPTHDDHYIRTDISEHYWLNGAkFIGTFF
IPDTYNPDDDKIYFFFRESSQEGSTSDKTILSRVGRVCKNDVGGQRSLINKWTTFLKARLIC
SIPGSDGADTYFDELQDIYLLPTRDERNPVVYGVFTTTSSIFKGSAVCVYSMADIRAVFNGP
YAHKESADHRWVQYDGRIPYPRPGTCPSKTYDPLIKSTRDFPDDVISFIKRHSVMYKSVYPV
AGGPTFKRINVDYRLTQIVVDHVIAEDGQYDVMFLGTDIGTVLKVVSIKSKWNMEEVLEE
LQIFKHSSIILNMELSLKQQQLYIGSRDGLVQLSLHRCDTYgKACADCCCLARDPYCAWDGNA
CSRYAPTSKRRARRQDVKYGDPITQCWDIEDSISHETADEKVIFGIEFNSTFLECIPKSQQA
TIKWYIQRSGDEHREELKPDERIIKTEYGLLIRSLQKKDSGMYyCKAQEHTFIHTIVKLTLN
VIENEQMENTQRAEHEEGQVKDLLAESRLRYKDYIQILSSPNFSLDQYCEQMWHREKRRQRN
KGGPKWKHMQEMKKKRNRHRDLDELPRAVAT

```

Important features of the protein:**Signal peptide:**

amino acids 1-36

N-glycosylation sites.

amino acids 139-142, 607-610, 724-727

Tyrosine kinase phosphorylation site.

amino acids 571-576

Gram-positive cocci surface proteins 'anchoring' hexapeptide.

amino acids 32-37

CCCTGACCTCCCTGAGCCACACTGAGCTGGAAGCCGCAGAGGTCATCCTGGAGCATGCCACCAGCGGGGAGCAGA
CAACCTCCCAGGTAAGCTGGGAGCAAGACCTGAAGCTGTTTCTTCAGGAGCCTGGTGTATTTTCCCCACCCCAC
CTCAGCAGTTCAGCCAGCAGGGACTGATCAGGTGTGTGCTCTGGAGTGGGAGCAGAAGGCGTGGCTGGCAAGA
GTGGCCTGGAGAAAGAGGTTTACGCGCTTGACCAGCCGAGCTGCCCGTGACTACAAGATCCAGAACCATGGGCATC
GGGTGAGGTGGGGGGGACAGGTGTCTATGTGCACCTTTCTGTCTCAGCAAGAAGAGCTGAGAGAGGGGATCTTGG
AGCCATTGAGGGTGTATGGAGCTACAGAGGGGAGGGAAAGGTATTTAAGTTAACAGTGTGGCACAATAGTTAA
GAGCACAGTTTTTGGAGCTAGACCGACATAGGTTCAAATTTCTCTTCTGTTGCTTCTAGTTCTGTAGCCCCAGGT
AAGGGAGTGACTTAACCTCTCTGGACTTCAATTTCTCATCTACTAAAGTAGGGCCAATAATAGCACCCACCTCAT
AGGGAAGATTAAATGACATAATGTATGTGATGCAACTAGCAAAGTACCAGTCCCATAGTAAGTCATGCCCCACA
TATTTCCACCCACCCCTGTTCTCTGCCTTCCCAACCAGGTAAGTCAACGACTGGAGCAGAGGCGGCAGCAGGCTT
CAGAGCGGGAGGCTCCAAGTACAGACAGAGTTTACAGGAAGTGCAGAGAGCATCCGCCGGGCACAGGTGAGCC
AGGTGAAGGGGGCTGCCCGCTGGCCCTGTCGAGGGGGCTGGCTTAGATGTGGAGCGCTGGCTGAAGCCAGCCA
TGACCCAGGCCCAGGATGAGGTGGAGCAGGAGCGGCGGCTCAGTGAGGTCTGGCTGCCAGAGGACCTCTCTC
CAACCGCTGAGGATGCTGAGCTTTCTGACTTTGAGGAATGTGAGGAGACGGGAGAGCTCTTTGAGGAGCCTGCC
CCCAAGCCCTGGCCACGAGGGCCCTCCCCTGCCCTGCACACGTGGTATTTGCTATCAGGCAGGGCGTGAGGATG
AGCTGACAATCACGGAGGGTGAGTGGCTGGAGGTATAGAGGAGGGAGATGCTGACGAATGGGTCAAGGCTCGGA
ACCAGCACGGCGAGGTAGGCTTTGTCCCTGAGCGATATCTCAACTTCCCGGACCTCTCCCTCCCAGAGAGCAGCC
AAGACAGTGACAATCCCTGCGGGGACGAGCCACAGCATCTCTGGCAGCGCCCTGTACAGCTACACGGACAGA
GTGCAGAGGAGCTGAGCTTCCCTGAGGGGGCACTCATCGTCTGCTGCCCGGGCCAAAGTAGGATGATGACG
GCTTCTGGAGGGGAGAATTTGGGGGCCGTGTTGGGGTCTTCCCCTCCCTGCTGGTGGAAGAGCTGCTTGGCCCC
CAGGGCCACCTGAACCTCTCTGACCCTGAACAGATGCTGCCGTCCCCTTCTCCTCCCAGCTTCTCCCCACCTGCAC
CTACCTCTGTGTTGGATGGGCCCCCTGCACCTGTCTGCCCTGGGACAAAGCCCTGGACTTCCCTGGGTTCCTGG
ACATGATGTGCACCTCGACTCAGGCCGATGCGTCCACCCACTCCCCCGCGGCTAAAGCCCCGGATCCTGGCCACC
CAGATCCCCTCACCTGAGAGCCAGGGAAGCTTACCCCCAGTGATGCTGCTGCTCCTATCTTCAAGCTGTGAGA
CCACACCATCAATGATCCAGAGCAACACAGCCAAAGCTGGAATCGCCCTATTTTCCACCTCACTCCAAGGGT
GGAAACTTGGCCCTTCCCATTTCTAGAGCTGGAACCCACTCCTTTTTTTTCCCATTTGTTCTATCATCTCTAGGACC
GGAATACTACCTTCTCTTCTGTCTATGACCCTATCTAGGGTGGTGAAATGCCTGAAATCTCTGGGGCTGGAAACC
ATCCATCAAGTCTCTAGTAGTTCTGGGCCACCTCTTTCCCACCCCTGGCTCCATGACCCACCCCACTCTGGATG
CCAGGTCACCTGGGCTTGGGCTGGGGAGGGAACAGGCCCTTGGGAATCAGGAGCTGGAGCCAGGATGCGAAGCAG
CTGTAATGGTCTGAGCGGATTTATTTGACAATGAATAAGGGCAGGAAGCCAGGCGAGGCGCTGGGCTCTGTG
CTAAGAGGGCAGGGGGCCTACGGTGCTATTGCTTTAGGGGCCACCAAGGCGAGGCGCTGCTCCAGCTGCCAC
GCTCTATCATATGGAGCGAGGTGTTGGGGAAGGCGGGGCAGGCAGCCTGTTGCAGGCAGGGGAAGGAGAAGAGAC
TGAGGGGGCTGTGACCTCTCCTGAGGCCCCAGCCTGAGAGCTGTGCAACTCCAGGTGGAAGTAGAGCTGGTCCCTC
AGCTGGGGGCGAGTGCTGTCCAGTGAGGGGAGGGGCTTTCACGCCCCACCCCACTCCCTGGCCCTGCCAGCTGGTAG
TCCATCAGCACAATGAAGGAGACTTGGAGAAGAGGAAGAATAACACTGTTGCTTCTGTTCAAGCTGTGTCCAGC
TTTTCCCCTGGGGCTCCAGGACCTTCCCTACCTCCACCACCAACCAAGGATTTATGACAAAGGCTAAGCCTGC
AGTTTACTCTGGGGGTTTCAAGGAGCCGAAAGGCTTAAATAGTTTAAAGTAGGTGATGGGAAGATGAGATTACCTCA
TTTAGGGCTCAGGCAGACTCACCTCACATACTCCCTGCTCCCTGTGGTAGAGACACCTGAGAGAAAGGGGAGGGG
TCAACAATGAGAGACCAGGAGTAGGTCTATCAGTGCCCCCAGAGTAGAGAGCAATAAGAGCCAGCCAGTGC
AGTCCCGGCTGTGTTTTCTACCTGGTGATCAGAAGTGCTGTGGTTTGTCTGGCTGCCCACTTGGCTCTTGGTGG
CGACCCCTGGGCTTGGGCCCCCTCCCGGCCCTCAGTGTGTGGCTCTGCAGAAGCTCTGGGCTTCCCTTCAAGT
CACGAGGGGTTAGGCTGCTGTCCCTGAGTCTCCCTCATTTGTAAGTGGGGGCTGGCTAGGAGCTGGGCTGTGGCC
TCTCAGGGGGCAGCCTCTCCATGGCAGGCATCCCTGCCTTGGGCTGCCCTCCCCAGACCCCTGACCACCCCTG
GGTCTGTCCCCCACCAGAGCCCAGCTCCTGTCTGTGGGGGAGCCATCACGGTGTTCGTGCAGTCCATAGCGCT
TCTCAATGTGTGTACCCCGAACCTGGGAGGGGAGGGAACACTGGGGTTTTAGGACCACAACCTCAGAGGCTGCTTG
GCCCTCCCCTCAGCAGGGGACATCTGAGTTTGGTGGCTACTTCCCTCTGGCCTAAGGTAGGGGAGGCCTTCTC
AGATTGTGGGACATTTGTGTAGCTGACTTCTGTGAGCTCCCAGTCCAGGAGGAAGAGCCAAAGGCCCACTT
TTGGGATCAGGTGCCTGATCACTGGGCCCCCTACCTCAGCCCCCTTTCCTGGAGCAGCTGCCCACTGCCCA
CAGAGAACACAGTGGTCTCCCTGTCCGGGGGCGGCTTTTTCTTCTTGGAGCGTCCCTGACGGACAAGTGAGG
GCCTCTTGTGTCGGCTGCAATGGATGCAAGGGGCTGCAGAGCCCAGGTGCACTGTGTGATGATGGGAGGGGGCTC
CGTCCCTGACGGCTGGAGGTGGCATCCACACTGGACAGCAGGAGGAGGGGAGTGAGGGTAACATTTCCATTTCCCT
TCATGTTTGGTTTCTTACGTTTCTTTCAGCATGCTCCTTAAACCCCAAGAGCCCCAATTTCCCCAAGCCCCATTT
TTTCTTGTCTTTATCTAATAAACTCAATATTAAG

FIGURE 178

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401
><subunit 1 of 1, 370 aa, 1 stop
><MW: 40685, pI: 4.53, NX(S/T): 0
MQLAKYQSHSKSCPTVFPPTPVLCLPNQVLQRLEQRRQQASEREAPSIEQRLQEVRESIRRA
QVSQVKGAARLALLQGAGLDVERWLKPAMTQAQDEVEQERRLSEARLSQRDLSPTAEDAELS
DFEECEETGELFEEPAPQALATRALPCPAHVVFYQAGREDELTITEGEWLEVIEEGDADEW
VKARNQHGEVGFVPERYLNFPDLSLPESSQSDSNPCGAEPATAFLAQALYSYTGQSAEELSFP
EGALIIRLLPRAQDGVDDGFWRGEFGGRVGVFPSLLVEELLGPPGPPPELSDPEQMLPSPSPPS
FSPPAPTSLVDGPPAPVLPDGDALDFPGFLDMMAPRLRPMRPPPPPPAKAPDPGHPDPLT
```


FIGURE 179A

CACAGGGAGACCCACAGACACATATGCACGAGAGAGACAGAGGAGGAAAGAGACAGAGACAAAGGCACAGCGGAA
GAAGGCAGAGACAGGGCAGGCACAGAAGCGGCCAGACAGAGTCTACAGAGGGAGAGGCCAGAGAAGCTGCAGA
AGACACAGGCAGGGAGAGACAAAGATCCAGGAAAGGAGGGCTCAGGAGGAGAGTTTGGAGAAGCCAGACCCCTGG
GCACCTCTCCCAAGCCCAAGGACTAAGTTTTCTCCATTTCTTTAACGGTCTCAGCCCTTCTGAAAACCTTTGCC
TCTGACCTTGGCAGGAGTCCAAGCCCCCAGGCTACAGAGAGGAGCTTTCCAAAGCTAGGGTGTGGAGGACTTGGT
GCCCTAGACGGCCTCAGTCCCTCCCAGCTGCAGTACCAGTGGC**ATGT**CCCAGACAGGCTCGCATCCCCGGGAGGGG
CTTGGCAGGGCGCTGGCTGTGGGGAGCCCAACCCTGCCTCCTGCTCCCCATTGTGCCGCTCTCCTGGCTGGTGTG
GCTGCTTCTGCTACTGCTGGCCTCTCTCCTGCCCTCAGCCCGGCTGGCCAGCCCCCTCCCCGGGAGGAGGAGAT
CGTGTTCAGAGAAGCTCAACGGCAGCGTCTGCCTGGCTCGGGCGCCCCCTGCCAGGCTGTTGTGCCGCTTGCA
GGCCTTTGGGGAGACGCTGCTACTAGAGCTGGAGCAGGACTCCGGTGTGCAGGTGAGGGGGCTGACAGTGCAGTA
CCTGGGCCAGGCGCTGAGCTGCTGGGTGGAGCAGAGCCTGGCACCTACCTGACTGGCACCATCAATGGAGATCC
GGAGTCGGTGGCATCTCTGCACTGGGATGGGGGAGCCCTGTTAGGCGTGTACAAATATCGGGGGGCTGAACCTCA
CCTCCAGCCCTGGAGGGAGGCACCCCTAACTCTGCTGGGGGACCTGGGGCTCACATCCTACGCCGGAAGAGTCC
TGCCAGCGGTCAAGGTCCCATGTGCAACGTCAAGGCTCCTCTTGGGAAGCCCCAGCCCCAGAGAGCCAA
GCGCTTTGCTTCACTGAGTAGATTTGTGGAGACACTGGTGGTGGCAGATGACAAGATGGCCGCTTCCACGGTGC
GGGGCTAAAGCGTACCTGCTAACAGTGATGGCAGCAGCAGCCAAAGGCCTTCAAGCACCCAAGCATCCGCAATCC
TGTCAGCTTGGTGGTGAAGTGGCTAGTATCCTGGGGTCAAGGCGAGGAGGGGGCCCCAAGTGGGGGCCAGTGTGC
CCAGACCTTGGCAGCTTCTGTGCCCTGGCAGCGGGGCTCAACACCCCTGAGGACTCGGGGCCCTGACCACCTTTGA
CACAGCCATTCTGTTACCCGTCAGGACCTGTGTGGAGTCTCCACTTGCGACACGCTGGGTATGGCTGATGTGGG
CACCTCTGTGACCCGGTTCGGAGCTGTGCCATTGTGGAGGATGATGGGCTCCAGTCAAGCCTTCACTGCTGTCTCA
TGAAGTGGGTGATGTCTTCAACATGCTCCATGACAACTCCAAGCCATGCATCAGTTTGAATGGGCTTTGAGCAC
CTCTCGCCATGTGATGGCCCTGTGATGGCTCATGTGGATCCTGAGGAGCCCTGGTCCCCCTGCAGTGGCCGCTT
CATCACTGACTTCCCTGGACAATGGCTATGGGCACTGTCTCTTAGACAAACCAGAGGCTCCATTGCATCTGCCTGT
GACTTTCCCTGGCAAGGACTATGATGCTGACCGCCAGTGCCAGCTGACCTTCGGGGCCGACTCACGCCATTGTCC
ACAGTGGCCGCGCCCTGTGCTGCCCTCTGGTGTCTGGCCACCTCAATGGCCATGCCATGTGCCAGACCAAACA
CTCGCCCTGGGCGGATGGCACACCCCTGCGGGGCCGACAGGCCCTGCATGGGTGGTGGTGGCTCCACATGGACCA
GCTCCAGGACTTCAATATTCCACAGGCTGGTGGCTGGGGTCTTGGGGACCATGGGGTGAAGTGTCTCGGACCTG
TGGGGGTGGTGTCCAGTTCTCCTCCCGAGACTGCACGAGGCTGTCCCCCGGAATGGTGGCAAGTACTGTGAGGG
CCGCCGTACCCGCTTCCGCTCCTGCAACACTGAGGACTGCCCAACTGGCTCAGCCCTGACCTTCCGCGAGGAGCA
GTGTGCTGCCTACAACCACCGCACCGACCTCTTCAAGAGCTTCCCAGGGCCCATGGACTGGGTTCCTCGCTACAC
AGGCGTGGCCCCCAGGACCAAGTGAACACTACCTGCCAGGCCCGGGCACTGGGCTACTACTATGTGCTGGAGCC
ACGGGTGGTAGATGGGACCCCTGTTCCCGGACAGCTCCTCGGTCTGTGTCCAGGGCCGATGCATCCATGTCTGG
CTGTGATCGCATCATTGGCTCCAAGAAGAAGTTTGACAAGTGCATGGTGTGCGGAGGGGACGGTTCTGGTTGACG
CAAGCAGTCAGGCTCCTTCAGGAAATTCAGGTACGGATACAACAATGTGGTCACTATCCCCCGGGGGGCCACCCA
CATTCTTGTCCGGCAGCAGGGAAACCTGGCCACCGGAGCATCTACTTGGCCCTGAAGCTGCCAGATGGCTCCTA
TGCCCTCAATGGTGAATACAGCTGATGCCCTCCCCACAGATGTGGTACTGCCTGGGGCAGTCAGCTTGCCTA
CAGCGGGGCCACTGCAGCCTCAGAGACACTGTGAGCCATGGGCCACTGGCCAGCCCTTTGACACTGCAAGTCTCT
AGTGGCTGGCAACCCCGAGGACACACGCCCTCCGATACAGCTTCTTCGTGCCCGGCCGACCCCTTCAACGCCACG
CCCCACTCCCCAGGACTGGCTGCACCGAAGAGCACAGATTCTGGAGATCCTTCGGCGGGCGCCCTGGGCGGGCAG
GAA**ATAA**ACCTCACTATCCCGGCTGCCCTTTCTGGGCACCGGGGCTCGGACTTAGCTGGGAGAAAGAGAGAGCTT
CTGTTGCTGCCTCATGCTAAGACTCAGTGGGGAGGGGCTGTGGGCGTGAGACCTGCCCTCCTCTGCCCTAAT
GCGCAGGCTGGCCCTGCCCTGGTTTCTGCCCTGGGAGGCAGTGTGGTGTAGTGGATGGAAGGGGCTGACAGAC
AGCCCTCGTAGATGGGACCCCTGTTCCCGGACAGCTCCTCGGTCTGTGTCCAGGGCCGATGCATCCATGTCTGG
CAGTTGTATTTTATTTAGTATTTTATTCACTTTTATTAGCACAGGGAAGGGGACAAGGACTAGGGTCTGGGGAA
CCTGACCCCTGACCCCTCATAGCCCTCACCTGGGGCTAGGAAATCCAGGGTGGTGGTATAGGTATAAGTGGTG
TGTGTATGCGTGTGTGTGTGTGTGTGAAAATGTGTGTGTGCTTATGTATGAGGTACAACCTGTTCTGCTTTCTC
TTCTGAATTTTATTTTGGGAAAAGAAAGTCAAGGGTAGGGTGGGCCCTCAGGGAGTGAGGGATTATCTTTT
TTTTTTTTTCTTTCTTTCTTTCTTTTGTGAGACAGAATCTCGCTCTGTGCCCCAGGCTGGAGTGCAATG
GCACAATCTCGGCTCACTGCATCCTCCGCTCCCGGTTCAAGTGATTCTCATGCCCTCAGCCTCCTGAGTAGCTG
GGATTACAGGCTCCTGCCACCACGCCAGCTAATTTTTGTGTTTTGTTTTGGAGACAGAGTCTCGCTATTGTG
ACCAGGGCTGGAATGATTTTCACTCACTGCAACCTTCGCCACCTGGGTCCAGCAATTTCTCCTGCCTCAGCCTCC
CGAGTAGCTGAGATTATAGGCACCTACCACCACGCCCGGCTAATTTTTGTATTTTGTAGTAGAGACGGGGTTTCAC
CATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTTAGGTGATCCACTCGCTTCATCTCCCAAAGTGCTGGGATT
ACAGGCGTGAGCCACCGTGCCTGGCCACGCCCACTAATTTTTGTATTTTGTAGTAGAGACAGGGTTTACCATTGT
TGCCAGGCTGCTCTTGAACCTCCTGAGTAACTGACCTGCCTCGCCCTCCCAAAGTGCTGGGATTACAGG
TGTGAGCCACCACGCCCGGTACATATTTTTTAAATTGAATTCTACTATTTATGTGATCCTTTTGGAGTCAGACAG

FIGURE 179B

ATGTGGTTGCATCCTAACTCCATGTCTCTGAGCATTAGATTTCTCATTTGCCAATAATAATACCTCCCTTAGAAG
TTTGTGTTGTGAGGATTAAATAATGTAAATAAAGAACTAGCATAACACTCAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAGGAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 180

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73492

><subunit 1 of 1, 837 aa, 1 stop

><MW: 90167, pI: 8.39, NX(S/T): 1

MSQTGSHPGRGLAGRWLWGAQPCLLLPIVPLSWLVWLLLLLLASLLPSARLASPLPREEEIV
 FPEKLNGSVLPGSGAPARLLCRLQAFGETLLLELEQDSGVQVEGLTVQYLGQAPELLGGAEP
 GTYLTGTINGDPESVASLHWDGGALLGVLQYRGAEHLQPLEGGTPNSAGGPGAHILRRKSP
 ASGQGPCMCNVKAPLGSPSPRRRAKRFASLSRFVETLVVADDKMAAFHGAGLKRYLLTVMAA
 AAKAFKHPSIRNPVSLVVTRLVILGSGEEGPQVGPSAAQTLRSFCAWQRLNTPEDSGPDHF
 DTAILFTRQDLCGVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHD
 NSKPCISLNGPLSTSRHVMAVMAHVDPEEPWSPCSARFITDFLDNGYGHCLLDKPEAPLHL
 PVTFFPGKDYDADRQCQLTFGPDSRHCPQLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCG
 PAQACMGGRCLHMDQLQDFNIPQAGGWGPWGPWGDCSRTC GGGVQFSSRDCTRVPVRNGGKY
 CEGRRTRFRSCNTEDCPTGSALTFREEQCAAYNHRTDLFKSFPGPMDWVPRYTGVAPQDQCK
 LTCQARALGYYYVLEPRVVDGTPCSPDSSSVCVQGRCIHAGCDRIIGSKKKFDKCMVCGGDG
 SGCSKQSGSFRKFRYGYNVVTIPAGATHILVRQQGNPGHRSIYLALKLPDGSYALNGEYTL
 MPSPTDVVLPGAVSLRYSGATAASETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVPRPT
 PSTPRPTQDWLHRRRAQILEILRRRPWAGRK

Important features of the protein:

Signal peptide:

amino acids 1-48

N-glycosylation site.

amino acids 68-71

Glycosaminoglycan attachment site

amino acids 188-191, 772-775

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 182-185

Tyrosine kinase phosphorylation site.

amino acids 730-736

N-myristoylation sites.

amino acids 5-10, 19-24, 121-126, 125-130, 130-135, 147-152, 167-
 172, 168-173, 174-179, 323-328, 352-357, 539-544, 555-560, 577-
 582, 679-684, 682-687, 763-768

Amidation sites.

amino acids 560-563, 834-837

Leucine zipper pattern.

amino acids 17-38, 24-45

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 358-367

FIGURE 181

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGG
 CAAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCC
 AAGAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCT
 AATTGTCCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACA
 TGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTG
 ACCAGAACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTT
 TAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGA
 TTAAAGTGATTCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACC
 ACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGA
 TTTTCTTAAAAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATC
 CCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTT
 CCTGCCAACGAAAAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGT
 AGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTG
 AAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGT
 CGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATA
 CTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTGTAAGTGGTGGGTGGCCC
 GCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAACTGCTGGCAACATAT
 AATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCTGGTAGCCAGCT
 CTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACC
 AAAAAAAAAAAAAAAAAA

FIGURE 182

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73727
><subunit 1 of 1, 317 aa, 1 stop
><MW: 37130, pI: 5.18, NX(S/T): 3
MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTTLIVLFWGSKHFWPEVPPKAY
DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKT
QIKVIPEFSEPEEEIDENEEITTTFFEQSVIWPAPKPIENRDFLKNKILEICDNVTMYWI
NPTLISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDY
TENGIEFDPMLDERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWV
ARMLGRV
```

Important features of the protein:**Signal peptide:**

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 183

GCGGAACCTGGCTCCGGCTGGCACCTGAGGAGCGGCGTGACCCCGAGGGCCCAGGGAGCTGCC
 CGGCTGGCCTAGGCAGGCAGCCGCACC**ATG**GCCAGCACGGCCGTGCAGCTTCTGGGCTTCCT
 GCTCAGCTTCCTGGGCATGGTGGGCACGTTGATCACCACCATCCTGCCGCACTGGCGGAGGA
 CAGCGCACGTGGGCACCAACATCCTCACGGCCGTGTCTACCTGAAAGGGCTCTGGATGGAG
 TGTGTGTGGCACAGCACAGGCATCTACCAGTGCCAGATCTACCGATCCCTGCTGGCGCTGCC
 CCAAGACCTCCAGGCTGCCCCGCGCCCTCATGGTCATCTCCTGCCTGCTCTCGGGCATAGCCT
 GCGCCTGCGCCGTTCATCGGGATGAAGTGACGCGCTGCGCCAAGGGCACACCCGCCAAGACC
 ACCTTTGCCATCCTCGGCGGCACCCTCTTCATCCTGGCCGGCCTCCTGTGCATGGTGGCCGT
 CTCCTGGACCACCAACGACGTGGTGCAGAACTTCTACAACCCGCTGCTGCCCAGCGGCATGA
 AGTTTGAGATTGGCCAGGCCCTGTACCTGGGCTTCATCTCCTCGTCCCTCTCGCTCATTGGT
 GGCACCCTGCTTTGCCTGTCCTGCCAGGACGAGGCACCCTACAGGCCCTACCAGGCCCCGCC
 CAGGGCCACCACGACCACTGCAAACACCGCACCTGCCTACCAGCCACCAGCTGCCTACAAAG
 ACAATCGGGCCCCCTCAGTGACCTCGGCCACGCACAGCGGGTACAGGCTGAACGACTACGTG
TGAGTCCCCACAGCCTGCTTCTCCCCTGGGCTGCTGTGGGCTGGGTCCCCGGCGGGACTGTC
 AATGGAGGCAGGGGTTCCAGCACAAAGTTTACTTCTGGGCAATTTTTGTATCCAAGGAAATA
 ATGTGAATGCGAGGAAATGTCTTTAGAGCACAGGGACAGAGGGGGAAATAAGAGGAGGAGAA
 AGCTCTCTATACCAAAGACTGAAAAAAAAAATCCTGTCTGTTTTTGTATTTATTATATATAT
 TTATGTGGGTGATTTGATAACAAGTTTAATATAAAGTGACTTGGGAGTTTGGTCAGTGGGGT
 TGGTTTGTGATCCAGGAATAAACCTTGCGGATGTGGCTGTTTATGAAAAAAAAAAAAA

FIGURE 184

MASTAVQLLGFLLSFLGMVGTLLITLPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIY
QCQIYRSLLALPQDLQAARALMVISCLLSGIACACAVIGMKCTRCAKGTPAKTTFAILGGTL
FILAGLLCMVAVSWTTNDVVQNFYNPLLPSPGMKFEIGQALYLGFISSSLIGGTLLCLSCQ
DEAPYRPHYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-103, 115-141, 160-182

FIGURE 185

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCG
CAGGCGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACCACATGCCAAGTGGTGGCGTTCCT
CCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCC
AGGACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGC
GTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGC
CATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCC
TGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCC
AACATGACACTGACCTCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGGAGT
GTCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCG
GCATGGGTGGGATGGTGCAGACTGTTACAGACCAGGTACACATTTGGTGCGGCTCTGTTTCGTG
GGCTGGGTGCGCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGG
CCTGGCACCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTC
CCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAG
ATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTA
TGTG**TAA**TGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAA
AACAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGC
CTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCC
ACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTAT
TTCTTTTTTTTAAATATAACTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCAC
ATTTTGATGATTTAGACAGACTCCCCCTCTTCCTCCTAGTCAATAAACCCATTGATGATCTA
TTTCCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTT
CTGCTGTTTGAATTTTGTCTCCCCACCCCCAACTTGGCTAGTAATAAACACTTACTGAAGAA
GAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTG
TGATCTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTG
CTGTTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAG
TCCTCTTTCTGTGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTTAAT
TTAAGTCCTAAATATAGTTAAAATAAATAATGTTTTAGTAAAATGATACACTATCTCTGTGA
AATAGCCTCACCCCTACATGTGGATAGAAGGAAATGAAAAAATAATTGCTTTGACATTGTCT
ATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCCATGAAAAGCTCACACCTGTAATC
CTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGTTCGAGACTAGCCTG
GGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGGTG
GCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGG
AGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGA
TCCTGTCTAAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAA
ACTAATTCTTTAA

FIGURE 186

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73734

><subunit 1 of 1, 261 aa, 1 stop

><MW: 27856, pI: 8.50, NX(S/T): 1

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTE
CRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIM
FIVSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTL
IGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTE
DEVQSYPSKHDYV

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

1001753-1001753

FIGURE 187

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCG
GAGTCCAGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTG
GGCTGTTTCTTGGTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGG
AGAGTGTGGCCTTCATTGAAAACAACATCGTGGTTTTTGAAGAACTTCTGGGAAGGACTGTG
GATGAATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCTTGCTGG
CTCTTTCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTC
TTGGCTTTCATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAA
GGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGC
TCATCCCTGTGAGCTGGGTGGCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAAT
GTTGCCCAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGTCT
GATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTACA
GATACTCGATACCTTCCCATCGCACAACCCAAAAAAGTTATCACACCGGAAAGAAGTCACCG
AGCGTCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGC
CATGCAAATGACAAAAATCTATATTACTTTCTCAAATGGACCCCAAAGAACTTTGATTTA
CTGTTCTTAAGTGCCTAATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAA
GCTATTTTCAGCAGAATGAGATATTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAAT
TTGTTTTCTAAGGTGGTTCAAGCATCTACTCTTTTATCATTTACTTCAAATGACATTGCT
AAAGACTGCATTATTTTACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAG
TGTAACATTTATATCTCACATAGAGACATGCTTATATGGTTTTATTTAAATGAAATGCCAG
TCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTG
AAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCATTTATAATGA
AGATTTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTG
TTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTT
TTCTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCA
AACTGCTTTTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGATG
GTTTTAGGAAAGTGAAAATATTTTTGTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAA
GAAATCATATATGTATGGATATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATC
AATATAAATAAAAGAGCAGAAAAATATGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACA
ACAAAAAAGTTGTCCTTTGAGAACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAATTG
TCATTTTTGTTCTGTGAAAAATAAATTTCTTCTTGTACCATTCTGTTTAGTTTTACTAAA
ATCTGTAAATACTGTATTTTTCTGTTTATCCAAATTTGATGAACTGACAATCCAATTTGA
AAGTTTGTGTGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTTTATACATTTATA
TTAATAAATTGTACATTTTTCTAATT

FIGURE 188

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73735

><subunit 1 of 1, 225 aa, 1 stop

><MW: 24845, pI: 9.07, NX(S/T): 0

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRM
QCKIYDSSLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGI
IFIITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVF
CCNEKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

1007254-24404

FIGURE 189

TCGCC**ATG**GCCTCTGCCGGAATGCAGATCCTGGGAGTCGTCTTGACACTGCTGGGCTGGGTG
AATGGCCTGGTCTCCTGTGCCCTGCCCATGTGGAAGGTGACCGCTTTCATCGGCAACAGCAT
CGTGGTGGCCCAGGTGGTGTGGGAGGGCCTGTGGATGTCCTGCGTGGTGCAGAGCACCGGCC
AGATGCAGTGCAAGGTGTACGACTCACTGCTGGCGCTGCCACAGGACCTGCAGGCTGCACGT
GCCCTCTGTGTCATCGCCCTCCTTGTGGCCCTGTTCTGGCTTGCTGGTCTACCTTGCTGGGGC
CAAGTGTACCACCTGTGTGGAGGAGAAGGATTCCAAGGCCCGCCTGGTGCTCACCTCTGGGA
TTGTCTTTGTTCATCTCAGGGGTCCTGACGCTAATCCCCGTGTGCTGGACGGCGCATGCCATC
ATCCGGGACTTCTATAACCCCCCTGGTGGCTGAGGCCCAAAGCGGGAGCTGGGGGCCCTCCCT
CTACTTGGGCTGGGCGGCCCTCAGGCCTTTTGTGCTGGGTGGGGGGTTGCTGTGCTGCACTT
GCCCCCTCGGGGGGGTCCCAGGGCCCCAGCCATTACATGGCCCCGCTACTCAACATCTGCCCCCT
GCCATCTCTCGGGGGGCCCTCTGAGTACCCTACCAAGAATTACGTCT**TGA**CGTGGAGGGGAATG
GGGGCTCCGCTGGCGCTAGAGCCATCCAGAAGTGGCAGTGCCCAACAGCTTTGGGATGGGTT
CGTACCTTTTGTCTCTGCCTCCTGCTATTTTTCTTTTGACTGAGGATATTTAAAATTCATTT
GAAAACTGAGCCAAGGTGTTGACTCAGACTCTCACTTAGGCTCTGCTGTTTCTCACCTTGG
ATGATGGAGCCAAAGAGGGGATGCTTTGAGATTCTGGATCTTGACATGCCCATCTTAGAAGC
CAGTCAAGCTATGGAATAATGCGGAGGCTGCTTGCTGTGCTGGCTTTGCAACAAGACAGAC
TGTCCTCCAAGAGTTCCTGCTGCTGCTGGGGGCTGGGCTTCCCTAGATGTCACTGGACAGCTG
CCCCCATCCTACTCAGGTCTCTGGAGCTCCTCTCTTACCCCCTGGAAAAACAAATCATCTG
TTAACAAAGGACTGCCACCTCCGGAACCTTCTGACCTCTGTTTCCCTCCGTCCTGATAAGACG
TCCACCCCCCAGGGCCAGGTCCCAGCTATGTAGACCCCCGCCCCACCTCCAACACTGCACC
CTTCTGCCCTGCCCCCTCGTCTCACCCCTTTTACACTCACATTTTTATCAAATAAAGCATG
TTTTGTTAGTGCA

FIGURE 190

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73736

><subunit 1 of 1, 220 aa, 1 stop

><MW: 23292, pI: 8.43, NX(S/T): 0

MASAGMQILGVVLTLLGWVNGLVSCALPMWKVTAFIGNSIVVAQVVWEGWLWMSCVVQSTGQM
QCKVYDSLLALPQDLQAARALCVIALLLVALFGLLVYLAGAKCTTCVEEKDSKARLVLTSGIV
FVISGVLTLIPVCWTAHAIIRDFYNPLVAEAQKREL GASLYLGWAASGLLLLGGGLLCCTCP
SGGSQGPHSHMARYSTSAPAISRGPSSEYPTKNYV

Transmembrane domains:

amino acids 8-30 (type II), 82-102, 121-140, 166-186

FIGURE 191

GCCAAGGAGAACATCATCAAAGACTTCTCTAGACTCAAAGGCTTCCACGTTCTACATCTTG
AGCATCTTCTACCACTCCGAATTGAACCAGTCTTCAAAGTAAAGGCAATGGCATTTTATCCC
TTGCAAATTGCTGGGCTGGTTCTTGGGTTCCCTTGGCATGGTGGGGACTCTTGCCACAACCCT
TCTGCCTCAGTGGTGGAGTATCAGCTTTTGTGTCAGCAACATTATTGTCTTTGAGAGGCTC
TGGGAAGGGCTCTGGATGAAATTGCATCCGACAAGCCAGGGTCCGGTTGCAATGCAAGTTCTA
TAGCTCCTTGTTGGCTCTCCCGCCTGCCCTGGAAACAGCCCGGGCCCTCATGTGTGTGGCTG
TTGCTCTCTCCTTGATCGCCCTGCTTATTGGCATCTGTGGCATGAAGCAGGTCCAGTGCACA
GGCTCTAACGAGAGGGGCCAAAGCATACCTTCTGGGAACTTCAGGAGTCCTCTTCATCCTGAC
GGGTATCTTCGTTCTGATTCCGGTGAGCTGGACAGCCAATATAATCATCAGAGATTTCTACA
ACCCAGCCATCCACATAGGTCAGAAACGAGAGCTGGGAGCAGCACTTTTCCTTGGCTGGGCA
AGCGCTGCTGTCCTCTTCATTGGAGGGGGTCTGCTTTGTGGATTTTGCTGCTGCAACAGAAA
GAAGCAAGGGTACAGATATCCAGTGCCTGGCTACCGTGTGCCACACACAGATAAGCGAAGAA
ATACGACAATGCTTAGTAAGACCTCCACCAGTTATGTCTTAATGCCTCCTTTTGGCTCCAAGT
ATGGACTATGGTCAATGTTTTTTATAAAGTCCTGCTAGAACTGTAAGTATGTGAGGCAGGA
GAAGTTGCTTTATGTCTAGATTTACATTGATACGAAAGTTTCAATTTGTTACTGGTGGTAGG
AATGAAAATGACTTACTTGGACATTCTGACTTCAGGTGTATTAAATGCATTGACTATTGTTG
GACCCAATCGCTGCTCCAATTTTCATATTCTAAATTCAAGTATACCCATAATCATTAGCAAG
TGTACAATGATGGACTACTTATTACTTTTTGACCATCATGTATTATCTGATAAGAATCTAAA
GTTGAAATTGATATTCTATAACAATAAAACATATACCTATTCTA

FIGURE 192

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73737

><subunit 1 of 1, 173 aa, 1 stop

><MW: 18938, pI: 9.99, NX(S/T): 1

MNCIRQARVRLQCKFYSSLLALPPALETARALMCVAVALSLLIALLIGICGMKQVQCTGSNER
AKAYLLGTSGVLFILTGIFVLIPVSWTANIIIRDFYNPAIHIGQKRELGAALFLGWASAAVL
FIGGGLLCGFCCCNRRKQGYRYPVPGYRVPHTDKRRNTTMLSSTSYV

Important features of the protein:

Transmembrane domains:

amino acids 31-51, 71-90, 112-133

N-glycosylation site.

amino acids 161-164

FIGURE 193

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGCATGAAAGATCACTGGGGGT
CTCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGTCTCC
AAAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCCATCA
CATACCTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACTTGTGTACC
GAGAGCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGCTAAATTCTCCA
TGGACATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTG
AGTTTCTTTCAGTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTCAGAGTAATCTTG
ACTGAATGGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCAT
TTTTTTTTTAACACGTCAATAAAAAAATAATCTCCCAGA

FIGURE 194

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73739

><subunit 1 of 1, 85 aa, 1 stop

><MW: 9232, pI: 7.94, NX(S/T): 0

MKITGGLLLLCTVVYFCSSSEAASLSPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITYGN
ECHLCTESLKSNGRVQFLHDGSC

Signal peptide:

amino acids 1-19

10073739

FIGURE 195

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCTCGCCCTCCTGCTCGCGCCC
CGCCGCC**ATG**GCTGCCTCCCCCGCGCGGCCTGCTGTCTGGCCCTGACCGGGCTGGCGCTGC
TCCTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAAATAAACTCAAGCTGATGCTTCAA
AAACGAGAAGCACCTGTTCCAATAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGA
ATTCCTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGGCCCCGAGGTGC
AGCAGTGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATC
ACCTATTGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCA
CTATGATGAAGACTCTGCAATTGGTCCCCGGAGCCCCCTACGGCTTTAGGCATGGAGCCAGCG
TCAACTACGATGACTAC**TAA**CCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATTTC
TCTTCATGTATCTCCTAATGCCTTACACTACTTGGTTTCTGATTTGCTCTATTTTCAGCAGAT
CTTTTCTACCTACTTTGTGTGATCAAAAAAGAAGAGTTAAACAACACATGTAAATGCCTTT
TGATATTTTCATGGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTAAAAAGA

FIGURE 196

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73742

><subunit 1 of 1, 148 aa, 1 stop

><MW: 17183, pI: 8.77, NX(S/T): 0

MAASPARPAVLALTGLALLLLLCWGPGGISGNKLLKMLQKREAPVPTKTKVAVDENKAKEFL
GSLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNDRDRNGHEYYGDDYQRHYD
EDSAIGPRSPYGFRHGASVNYDDY

Signal peptide:

amino acids 1-30

199/249

CGGCTCGAGCCCGCCGGAAGTGCCCGAGGGGGCGCGATGGAGCTGGGGGAGCCGGGCGCTC
GGTAGCGCGGCGGGCAAGGCAGGCGCCATGACCCTGATTGAAGGGGTGGGTGATGAGGTGAC
CGTCCTTTTCTCGGTGCTTGCCTGCCTTCTGGTGCTGGCCCTTGCCTGGGTCTCAACGCACA
CCGCTGAGGGCGGGGACCCACTGCCCCAGCCGTCAGGGACCCCAACGCCATCCCAGCCCAGC
GCAGCCATGGCAGCTACCGACAGCATGAGAGGGGAGGCCCCAGGGGCAGAGACCCCCAGCCT
GAGACACAGAGGTCAAGCTGCACAGCCAGAGCCCAGCACGGGGTTCACAGCAACACCGCCAG
CCCCGGACTCCCCGCAGGAGCCCCTCGTGCTACGGCTGAAATTCTCAATGATTCAGAGCAG
GTGGCCAGGGCCTGGCCCCACGACACCATTGGCTCCTTGAAAAGGACCCAGTTTCCCGGCCG
GGAACAGCAGGTGCGACTCATCTACCAAGGGCAGCTGCTAGGCGACGACACCAGACCCTGG
GCAGCCTTCACCTCCCTCCCAACTGCGTTCTCCACTGCCACGTGTCCACGAGAGTCGGTCCC
CCAAATCCCCCCTGCCCGCCGGGGTCCGAGCCCGGCCCTCCGGGCTGGAAATCGGCAGCCT
GCTGCTGCCCCTGCTGCTCCTGCTGTTGCTGCTGCTCTGGTACTGCCAGATCCAGTACCGGC
CCTTCTTTCCCCTGACCGCCACTCTGGGCCTGGCCGGCTTCACCCTGCTCCTCAGTCTCCTG
GCCTTTGCCATGTACCGCCCGTTAGTGCCTCCGCGGGCGCTTGGCAGCGTCGCCGGCCCCCTCC
GGACCTTGCTCCCCGCGCCGCGGGGAGCTGCTGCCTGCCAGGCCCGCCTCTCCGGCCTG
CCTCTTCCCGCTGCCCTGGAGCCCAGCCCTGCGCCGCAGAGGACTCCCGGGACTGGCGGAGG
CCCCGCCCTGCGACCGCCGGGGCTCGGGGCCACCTCCCGGGGCTGCTGAACCTCAGCCCGCA
CTGGGAGTGGGCTCCTCGGGGTCGGGCATCTGCTGTCGCTGCCTCGGCCCCGGGCAGAGCCG
GGCCGCCCCGGGGGCCCGTCTTAGTGTTCTGCCGGAGGACCCAGCCGCCTCCAATCCCTGAC
AGCTCCTTGGGCTGAGTTGGGGACGCCAGGTCGGTGGGAGGCTGGTGAAGGGGAGCGGGGAG
GGGCAGAGGAGTTCCCCGGAACCCGTGCAGATTAAAGTAACTGTGAAGTTTTAAAAA
AAAAA

FIGURE 198

MTLIEGVGDEVTVLFSVLACLLVLALAWVSTHTAEGGDPLPQPSGTPTPSQPSAAMAATDSM
RGEAPGAETPSLRHRGQAAQPEPSTGFTATPPAPDSPQEPLVLRLKFLNDSEQVARAWPHDT
IGSLKRTQFPGREQQVRLIYQGQLLGDDTQTLGSLHLPPNCVLHCHVSTRVGPPNPPCPPGS
EPGPSGLEIGSLLLPLLLLLLLLLLWYCQIQYRPFFPLTATLGLAGFTLLLSLLAFAMYRP

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 195-217

bioRxiv preprint doi: <https://doi.org/10.1101/001111>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 199

GAGATTGGAAACAGCCAGGTTGGAGCAGTGAGTGAGTAAGGAAACCTGGCTGCCCTCTCCAG
ATTCCCCAGGCTCTCAGAGAAGATCAGCAGAAAGTCTGCAAGACCCTAAGAACCATCAGCCC
TCAGCTGCACCTCCTCCCCTCCAAGGATGACAAAGGCGCTACTCATCTATTTGGTCAGCAGC
TTTCTTGCCCTAAATCAGGCCAGCCTCATCAGTCGCTGTGACTTGCCCCAGGTGCTGCAGCT
GGAGGACTTGGATGGGTTTGAGGGTTACTCCCTGAGTGACTGGCTGTGCCTGGCTTTTGTGG
AAAGCAAGTTCAACATATCAAAGATAAATGAAAATGCGGATGGAAGCTTTGACTATGGCCTC
TTCCAGATCAACAGCCACTACTGGTGCAACGATTATAAGAGTTACTCGGAAAACCTTTGCCA
CGTAGACTGTCAAGATCTGCTGAATCCCAACCTTCTTGACGGCATCCACTGCGCAAAAAGGA
TTGTGTCCGGAGCACGGGGGATGAACAACCTGGGTAGAATGGAGGTTGCACTGTTCAGGCCGG
CCACTCTCCTACTGGCTGACAGGATGCCGCCTGAGATTGAAACAGGGTGCGGGTGCACCGTGG
AGTCATTCCAAGACTCCTGTCTCACTCAGGGATTCTTCATTTCTTCTTCTCCTACTGCCTCCA
CTTCATGTTATTTTCTTCCCTTCCCATTACAACTAAAACCTGACCAGAGCCCCAGGAATAAA
TGGTTTTCTTGGCTTCCTCCTTACTCCCATCTGGACCCAGTCCCCTGGTTCCTGTCTGTTAT
TTGTAAACTGAGGACCACAATAAAGAAATCTTTATATTTATCG

FIGURE 200

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73746

><subunit 1 of 1, 148 aa, 1 stop

><MW: 16896, pI: 6.05, NX(S/T): 1

MTKALLIYLVSSFLALNQASLISRCDLAQVLQLEDLDGFEGYSLSDWLCLAFVESKFNISKI
NENADGSFDYGLFQINSHYWCNDYKSYSENLCNVDCQDLLNPNLLAGIHCAKRIVSGARGMN
NWVEWRLHCSGRPLSYWLTGCRLR

Signal peptide:

amino acids 1-18

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TCTGACCTGACTGGAAGCGTCCAAAGAGGGGACGGCTGTGACCCCTGCTTGACTGAGAACCCCA
 CCAGCTCATCCCAGACACCTCATAGCAACCTATTTATACAAAGGGGGAAAGAAACACCTGAG
 CAGAATGGAATCATTATTTTTTTTCCCAAGGAGAAAACCGGGGTAAAGGGAGGGAAGCAATTC
 AATTTGAAGTCCCTGTGAATGGGCTTTTCTGATTGTAAGTGAAGCAGGTCTTGACACACG
 TTGGGGTGAAACTTGGGTCCTGTGGTTTTCTGATTGTAAGTGAAGCAGGTCTTGACACACG
 TGTGGCAAATGTCAGGACCAGGTTAAGTGAAGTGGCAGAAAACTTCCAGGTGGAACAAGCA
 ACCCATGTTCTGCTGCAAGCTTGAAGGAGCCTGGAGCGGGAGAAAGCTAACTTGAACATGAC
 CTGTTGCATTTGGCAAGTTCTAGCAAC**ATG**CTCCTAAGGAAGCGATACAGGCACAGACCATG
 CAGACTCCAGTTCTCTGCTGCTCCTGATGCTGGGATGCGTCCTGATGATGGTGGCGATGT
 TGCACCCTCCCCACCACACCCTGCACCAGACTGTCACAGCCCAAGCCAGCAAGCACAGCCCT
 GAAGCCAGGTACCGCCTGGACTTTGGGGAATCCCAGGATTGGGTACTGGAAGCTGAGGATGA
 GGGTGAAGAGTACAGCCCTCTGGAGGGCCTGCCACCCTTTATCTCACTGCGGGAGGATCAGC
 TGCTGGTGGCCGTGGCCTTACCCAGGCCAGAAGGAACCAGAGCCAGGGCAGGAGAGGTGGG
 AGCTACCGCCTCATCAAGCAGCCAAGGAGGCAGGATAAGGAAGCCCCAAAGAGGGACTGGGG
 GGCTGATGAGGACGGGGAGGTGTCTGAAGAAGAGGAGTTGACCCCGTTCAGCCTGGACCCAC
 GTGGCCTCCAGGAGGCACTCAGTGCCCGCATCCCCCTCCAGAGGGCTCTGCCCGAGGTGCGG
 CACCCACTGTGTCTGCAGCAGCACCCTCAGGACAGCCTGCCACAGCCAGCGTCATCCTCTG
 TTTCCATGATGAGGCCTGGTCCACTCTCCTGCGGACTGTACACAGCATCCTCGACACAGTGC
 CCAGGGCCTTCTGAAGGAGATCATCCTCGTGGACGACCTCAGCCAGCAAGGACAACCTCAAG
 TCTGCTCTCAGCGAATATGTGGCCAGGCTGGAGGGGGTGAAGTTACTCAGGAGCAACAAGAG
 GCTGGGTGCCATCAGGGCCCCGGATGCTGGGGGCCACCAGAGCCACCGGGGATGTGCTCGTCT
 TCATGGATGCCCACTGCGAGTGCCACCAGGCTGGCTGGAGCCCCTCCTCAGCAGAATAGCT
 GGTGACAGGAGCCGAGTGGTATCTCCGGTGATAGATGTGATTGACTGGAAGACTTTCCAGTA
 TTACCCCTCAAAGGACCTGCAGCGTGGGGTGTTGGACTGGAAGCTGGATTTCCACTGGGAAC
 CTTTGCCAGAGCATGTGAGGAAGGCCCTCCAGTCCCCCATAAGCCCCATCAGGAGCCCTGTG
 GTGCCCGGAGAGGTGGTGGCCATGGACAGACATTACTTCCAAAACACTGGAGCGTATGACTC
 TCTTATGTGCTGCGAGGTGGTGAAAACCTCGAACTGTCTTTCAAGGCCTGGCTCTGTGGTG
 GCTCTGTTGAAATCCTTCCCTGCTCTCGGGTAGGACACATCTACCAAATCAGGATTCCCAT
 TCCCCCCTCGACCAGGAGGCCACCCTGAGGAACAGGGTTTCGCATTGCTGAGACCTGGCTGGG
 GTCATTCAAAGAAACCTTCTACAAGCATAGCCCAGAGGCCTTCTCCTTGAGCAAGGCTGAGA
 AGCCAGACTGCATGGAACGCTTGCAGCTGCAAAGGAGACTGGGTGTGCGGACATTCCACTGG
 TTTCTGGCTAATGTCTACCCTGAGCTGTACCCATCTGAACCCAGGCCAGTTTCTCTGGA
 GCTCCACAACACTGGACTTGGGCTCTGTGCAGACTGCCAGGCAGAAGGGGACATCCTGGGCT
 GTCCCATGGTGTGGCTCCTTGCAAGTGCAGCCGGCAGCAACAGTACCTGCAGCACACCAGC
 AGGAAGGAGATTCACTTTGGCAGCCACAGCACCTGTGCTTTGCTGTGAGGCAGGAGCAGGT
 GATTCTTCAGAACTGCACGGAGGAAGGCCTGGCCATCCACCAGCAGCACTGGGACTTCCAGG
 AGAATGGGATGATTGTCCACATTCTTTCTGGGAAATGCATGGAAGCTGTGGTGCAAGAAAAC
 AATAAAGATTTGTACCTGCGTCCGTGTGATGGAAAAGCCCCGCCAGCAGTGGCGATTTGACCA
 GATAAATGCTGTGGATGAACGA**TGA**ATGTCAATGTCAGAAGGAAAAGAGAATTTTGGCCATC
 AAAATCCAGCTCCAAGTGAACGTAAAGAGCTTATATATTTTCATGAAGCTGATCCTTTTGTGT
 GTGTGCTCCTTGTGTTAGGAGAGAAAAAGCTCTATGAAAGAATATAGGAAGTTTCTCCTTT
 TCACACCTTATTTCAATTGACTGCTGGCTGCTTA

FIGURE 202

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73760

><subunit 1 of 1, 639 aa, 1 stop

><MW: 73063, pI: 6.84, NX(S/T): 2

MLLRKRYRHRPCRLQFLLLLLMLGCVLMMVAMLHPPHHTLHQTVTAQASKHSPEARYRLDFG
ESQDWVLEAEDEGEYEYSPLEGLPPFISLREDQLLVAVALPQARRNQSQGRRGGSYRLIKQPR
RQDKEAPKRDWGADEDEGEVSEEEELTPFSLDPRGLQEALSARIPLQRALPEVRHPLCLQQHP
QDSLPTASVILCFHDEAWSTLLRTVHSILDTVPRAFLKEIILVDDLSQQGQLKSALSEYVAR
LEGVKLLRSNKRLGAIRARMLGATRATGDVLVFMDAHCECHPGWLEPLLSRIAGDRSRVVSP
VIDVIDWKTFQYYPSKDLQRGVLDWKLDHFHWEPLPEHVRKALQSPISPIRSPVVPGEVVAMD
RHYFQNTGAYDSLMSLRGGENLELSFKAWLCGGSVEILPCSRVGHIYQNQDSHSPLDQEATL
RNRVRIAETWLGSFKETFYKHSPEAFSLSKAEKPDCMERLQLQRRLGCRTFHWFLANVYPEL
YPSEPRPSFSGKLHNTGLGLCADCQAEGLGCPMVLAPCSDSRQQQYLQHTSRKEIHFGSP
QHLCFAVRQEQVILQNCTEEGLAIHQHWDFFQENGMIVHILSGKCMEEAVVQENNKDLYLRPC
DGKARQQWRFDQINAVDER

Signal peptide:

amino acids 1-28

FIGURE 203

CGCCAAGCATGCAGTAAAGGCTGAAAATCTGGGTACACAGCTGAGGAAGACCTCAGACATGGA
GTCCAGGATGTGGCTGCGCTGCTGCTGTCCCACCTCCTCCCTCTCTGGCCACTGCTGTTGC
TGCCCCCTCCCACCGCCTGCTCAGGGCTCTTCATCCTCCCCTCGAACCCACACAGCCCCAGCC
CGCCCCCGTGTGCCAGGGGAGGGCCCCCTCGGCCCCACGTCATGTGTGCGTGTGGGAGCGAGC
ACCTCCACCAAGCCGATCTCCTCGGGTCCCCAAGATCACGTCGGCAAGTCCTGCCTGGCACTG
CACCCCCAGCCACCCCATCAGGCTTTGAGGAGGGGGCCGCCCTCATCCCAATACCCCTGGGCT
ATCGTGTGGGGTCCCACCGTGTCTCGAGAGGATGGAGGGGACCCCAACTCTGCCAATCCCGG
ATTTCTGGACTATGTTTTTGAGCCCCCTCATGGGCTCGCAACCCACACCCCAACTCAGACT
CCATGCGAGGTGATGGAGATGGGCTTATCCTTGAGAGGGCACCTGCCACCCTGCGGCCATTCT
CTGTTGCGGGGCGCTGGGGAAGGTGTGGACCCCCAGCTCTATGTACAATTACCATCTCCAT
CATCATTGTTCTCGTGGCCACTGGCATCATCTTCAAGTTCTGCTGGGACCGCAGCCAGAAGC
GACGCAGACCCTCAGGGCAGCAAGGTGCCCTGAGGCAGGAGGAGAGCCAGCAGCCACTGACA
GACCTGTCCCCGGCTGGAGTCACTGTGCTGGGGGCCCTTCGGGGACTCACCTACCCCCACCCC
TGACCATGAGGAGCCCCGAGGGGGACCCCGGCCCTGGGATGCCCCACCCCAAGGGGGCTCCAG
CCTTCCAGTTGAACCGGTGAGGGGCAGGGGCAATGGGATGGGAGGGCAAAGAGGGGAAGGCAAC
TTAGGTCTTCAGAGCTGGGGTGGGGGTGCCCTCTGGATGGGTAGTGAGGAGGCAGGCGTGGC
CTCCACAGCCCCCTGGCCCTCCCAAGGGGGCTGGACCAGCTCCTCTCTGGGAGGCACCCTTC
CTTCTCCCAGTCTCTCAGGATCTGTGTCCTATTCTCTGCTGCCATAACTCCAACTCTGCC
TCTTTGGTTTTTTTTCTCATGCCACCTTGTCTAAGACAACCTCTGCCCTCTTAACCTTGATTCCC
CCTCTTTGTCTTTGAACCTTCCCCTTCTATTCTGGCCTACCCCTTGTTTCTGACTGTGCCCTT
TCCCTCTTCTCTCAGGATTTCCCTTGGTGAATCTGTGATGCCCCAATGTTGGGGTGCAGCC
AAGCAGGAGGCCAAGGGGGCCGGCACAGCCCCCATCCCACTGAGGGTGGGGCAGCTGTGGGGA
GCTGGGGCCACAGGGGCTCCTGGCTCCTGCCCTTGACACACCACCGGAACACTCCCCAGCC
CCACGGGCAATCCTATCTGCTCGCCCTCCTGCAGGTGGGGGCCCTCACATATCTGTGACTTCG
GGTCCCTGTCCCCACCCTTGTGCACTCACATGAAAGCCTTGACACACTCACCTCCACCTTCAC
AGGCCATTTGCACACGCTCCTGCACCCTCTCCCCGTCCATACCGCTCCGCTCAGCTGACTCT
CATGTTCTCTCGTCTCACATTTGCACTCTCTCCTTCCCACATTCTGTGCTCAGCTCACTCAG
TGGTCAGCGTTTCTGACACACTTTACCTCTCATGTGCGTTTTCCCGGCCCTGATGTTGTGGTGG
TGTGCGGCGTGTCACTCTCTCCCTCATGAACACCCACCCACCTCGTTTTCCGCAGCCCCCTGC
GTGCTGCTCCAGAGGTGGGTGGGAGGTGAGCTGGGGGCTCCTTGGGGCCCTCATCGGTGATGG
TCTCGTCCCATTTCCACACCATTGTTTTCTGTCTCCCCATCCTACTCCAAGGATGCCGGCA
TCAACCTGAGGGCTCCCCCTTGGGAATGGGGTAGTGAGGCCCCAGACTTCACCCCCAGCCCA
CTGCTAAAATCTGTTTTCTGACAGATGGGTTTTGGGGAGTCGCCTGCTGCACTACATGAGAA
AGGGACTCCCATTTGCCCTTCCCTTTCTCCTACAGTCCCTTTTGTCTTGTCTGTCTGGCTG
TCTGTGTGTGTGCCATTCTCTGGACTTCAGAGCCCCCTGAGCCAGTCTCCTTCCAGCCT
CCCTTTGGGCCTCCCTAACTCCACCTAGGCTGCCAGGGACCGGAGTCAGCTGGTTCAAGGCC
ATCGGGAGCTCTGCCTCCAAGTCTACCCCTCCCTTCCCGGACTCCCTCCTGTCCCCTCCTTT
CCTCCCTCCTTCCCTCCACTCTCCTTCCCTTTTGCTTCCCTGCCCTTTCCCCCTCCTCAGGTT
CTTCCCTCCTTCTCACTGGTTTTTTCCACCTTCCCTCCTTCCCTTCTTCCCTGGCTCCTAGGCT
GTGATATATATTTTTGTATTATCTCTTTCTTCTTCTTGTGGTGATCATCTTGAATTACTGTG
GGATGTAAGTTTTCAAATTTTCAAATAAAGCCTTTGCAAGATAA

FIGURE 204

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393

><subunit 1 of 1, 243 aa, 1 stop

><MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGV
PGRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIIYLDQGSPEMN
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 205

GTTAACCAGCGCAGTCCTCCGTGCGTCCCGCCCGCGCTGCCCTCACTCCCGGCCAGG**ATGG**
CATCCTGTCTGGCCCTGCGCATGGCGCTGCTGCTGGTCTCCGGGGTTCTGGCCCCTGCGGTG
CTCACAGACGATGTTCCACAGGAGCCCGTGCCACGCTGTGGAACGAGCCGGCCGAGCTGCC
GTCGGGAGAAGGCCCCGTGGAGAGCACCAGCCCCGGCCGGGAGCCCGTGGACACCGGTCCCC
CAGCCCCCACCCTCGCGCCAGGACCCGAGGACAGCACC GCGCAGGAGCGGCTGGACCAGGGC
GGCGGGTCGCTGGGGCCCGGCGCTATCGCGGCCATCGTGATCGCCGCCCTGCTGGCCACCTG
CGTGGTGCTGGCGCTCGTGGTCGTCGCGCTGAGAAAGTTTTCTGCCTCC**TGA**AGCGAATAAA
GGGGCCGCGCCCGGCCGCGGCGCGACTCGGCAAAAAAAAAAAAAA

FIGURE 206

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76398
><subunit 1 of 1, 121 aa, 1 stop
><MW: 12073, pI: 4.11, NX(S/T): 0
MASCLALRMALLLVSGVLAPAVLTDDVPQEPVPTLWNEPAELPSGEGPVESTSPGREPVDTG
PPAPTVPAGPEDSTAQERLDQGGGSLGPGAIAAIVIAALLATCVVLALVVVALRKFSAS

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 91-110

Glycosaminoglycan attachment site.

amino acids 44-47

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 116-119

N-myristoylation site.

amino acids 91-96

FIGURE 207

GGCCGTTGGTTGGTGC GCGGCTGAAGGGTGTGGCGCGAGCAGCGTCGTTGGTTGGCCGGCGG
CGGGCCGGGACGGGC**ATG**GCCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGGCGCTGGCCCA
CGGCTGTCTGCACTGCCACAGCAACTTCTCCAAGAAGTTCTCCTTCTACCGCCACCATGTGA
ACTTCAAGTCCTGGTGGGTGGGCGACATCCCCGTGTCAGGGGCGCTGCTCACCGACTGGAGC
GACGACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGGAGAAGCTGGA
CCAAGTGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGAAGATGTACTTCC
CCGGGTATTTCCCCAACGAGCTGCGAAACATCTTCCGGGAGCAGGTGCACCTCATCCAGAAC
GCCATCATCGAAAGGCACCTGGCACCAGGCAGCTGGGGAGGAGGGCAGCTCTCCAGGGAGGG
ACCCAGCCTAGCACCTGAAGGATCAATGCCATCACCCCGCGGGGACCTCCCC**TAA**GTAGCCC
CCAGAGGCGCTGGGAGTGTTGCCACCGCCCTCCCCTGAAGTTTGCTCCATCTCACGCTGGGG
GTCAACCTGGGGACCCCTTCCCTCCGGGCCATGGACACACATACATGAAAACCAGGCCGCAT
CGACTGTCAGCACCGCTGTGGCATCTTCCAGTACGAGACCATCTCCTGCAACAACCTGCACAG
ACTCGCACGTGCGCTGCTTTGGCTATAACTGCGAGTAGGGCTCAGGCATCACACCCACCCGT
GCCAGGGCCCTACTGTCCCTGGGGTCCCAGGCTCTCCTTGGAGGGGGCTCCCCGCCTTCCAC
CTGGCTGTCATCGGGTAGGGCGGGGCCGTGGGTTAGGGGGCGCACCACTTCCAAGCCTGTGT
CCCACAGGTCCTCGGCGCAGTGGAAGTCAGCTGTCCAGGGCCTCCTGAACTACATAAATAAC
TGGCACAAGTAAGTCCCCTCCTCAAACCAACACAGGCAGTGTGTGTATGTGAGCACCTCGTG
GGTGAGTATGTGTGGGGCACAGGCTGGCTCCCTCAGCTCCACGTCCTAGAGGGGCTCCCGA
GGAGGTGGAACCTCAACCCAGCTCTGCGCAGGAGGCGGCTGCAGTCCTTTTCTCCCTCAAAG
GTCTCCGACCCTCAGCTGGAGGCGGGCATCTTTCCTAAAGGGTCCCCATAGGGTCTGGTTCC
ACCCCATCCCAGGTCTGTGGTCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCC
ATGGAGGGGCTGACTGCCCCACATTGCCTTTCAGACAGGACACGAGCATGAGGTAAGGCCGC
CCTGACCTGGACTTCAGGGGGAGGGGGTAAAGGGAGAGAGGAGGGGGGCTAGGGGGTCTCT
AGATCAGTGGGGGCACTGCAGGTGGGGCTCTCCCTATACCTGGGACACCTGCTGGATGTCAC
CTCTGCAACCACACCCATGTGGTGGTTTCATGAACAGACCACGCTCCTCTGCCTTCTCCTGG
CCTGGGACACACAGAGCCACCCCGGCCTTGTGAGTGACCCAGAGAAGGGAGGCCTCGGGAGA
AGGGGTGCTCGTAAGCCAACACCAGCGTGCCGCGGCCTGCACACCCTTCGGACATCCCAGGC
ACGAGGGTGTCGTGGATGTGGCCACACATAGGACCACACGTCCCAGCTGGGAGGAGAGGCCT
GGGGCCCCCAGGGAGGGAGGCAGGGGGTGGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCC
CCGCAGCCTGGTATCGCCAGCCTTAAGGTGTCTGGAGCCCCCACACTTGGCCAACCTGACCT
TGGAAGATGCTGCTGAGTGTCTCAAGCAGCACTGACAGCAGCTGGGCCTGCCCCAGGGCAAC
GTGGGGGCGGAGACTCAGCTGGACAGCCCCTGCCTGTCACTCTGGAGCTGGGCTGCTGCTGC
CTCAGGACCCCCCTCTCCGACCCCGGACAGAGCTGAGCTGGCCAGGGCCAGGAGGGCGGGAGG
GAGGGAATGGGGGTGGGCTGTGCGCAGCATCAGCGCCTGGGCAGGTCCGCAGAGCTGCGGGA
TGTGATTAAAGTCCCTGATGTTTCTC

FIGURE 208

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76399
```

><subunit 1 of 1, 157 aa, 1 stop

><MW: 17681, pI: 7.65, NX(S/T): 1

MAALLCLVCLTAALAHGCLHCHSNFSSKKFSFYRHHVNFKSWWVGDI PVSGALLTDWSDDTMK

ELHLAI PAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNAIIER

HLAPGSWGGGQLSREGPSLAPEGSMPSPRGDLP

Signal peptide:

amino acids 1-15

FIGURE 209

AGCAGGAGCAGGAGAGGGACA**ATG**GAAGCTGCCCCGTCCAGGTTTCATGTTCCCTCTTATTTCT
 CCTCACGTGTGAGCTGGCTGCAGAAGTTGCTGCAGAAGTTGAGAAATCCTCAGATGGTCCTG
 GTGCTGCCCAGGAACCCACGTGGCTCACAGATGTCCCAGCTGCCATGGAATTCATTGCTGCC
 ACTGAGGTGGCTGTCATAGGCTTCTTCCAGGATTTAGAAATACCAGCAGTGCCCATACTCCA
 TAGCATGGTGCAAAAATTCCCAGGCGTGTCAATTTGGGATCAGCACTGATTCTGAGGTTCTGA
 CACACTACAACATCACTGGGAACACCATCTGCCTCTTTTCGCCTGGTAGACAATGAACAACCTG
 AATTTAGAGGACGAAGACATTGAAAGCATTGATGCCACCAAATTGAGCCGTTTCATTGAGAT
 CAACAGCCTCCACATGGTGACAGAGTACAACCCTGTGACTGTGATTGGGTTATTCAACAGCG
 TAATTCAGATTCATCTCCTCCTGATAATGAACAAGGCCTCCCCAGAGTATGAAGAGAACATG
 CACAGATACCAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAG
 TGGTATGAAAGAAAATGGGAAGGTGATATCATTTTTTCAAACATAAGGAGTCTCAACTGCCAG
 CTTTGGCAATTTACCAGACTCTAGATGACGAGTGGGATACACTGCCCACAGCAGAAGTTTCC
 GTAGAGCATGTGCAAACTTTTGTGATGGATTCCCTAAGTGGAAAATTGTTGAAAGAAAATCG
 TGAATCAGAAGGAAAGACTCCAAAGGTGGAACCTC**TGA**CTTCTCCTTGGAACACATATGGCC
 AAGTATCTACTTTTATGCAAAGTAAAAAGGCACAACCTCAAATCTCAGAGACACTAAACAACAG
 GATCACTAGGCCTGCCAACCACACACACACGCACGTGCACACACGCACGCACGCGTGCACAC
 ACACACGCGCACACACACACACACACAGAGCTTCATTTCTGTCTTAAAATCTCGTTTTCTC
 TTCTTCCTTCTTTTAAATTTTCATATCCTCACTCCCTATCCAATTTCTTCTTATCGTGCATT
 CATACTCTGTAAGCCCATCTGTAACACACCTAGATCAAGGCTTTAAGAGACTCACTGTGATG
 CCTCTATGAAAGAGAGGCATTCCTAGAGAAAGATTGTTCCAATTTGTCATTTAATATCAAGT
 TTGTATACTGCACATGACTTACACACAACATAGTTCCCTGCTCTTTTAAGGTTACCTAAGGGT
 TGAAACTCTACCTTCTTTCATAAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAAGG
 ATGGTTTTAAACACCTTTGTGAAATTGTCTTTTTGCCAGAAGTTAAAGGCTGTCTCCAAGTC
 CCTGAACTCAGCAGAAATAGACCATGTGAAACTCCATGCTTGGTTAGCATCTCCAACCTCCC
 TATGTAAATCAACAACCTGCATAATAAATAAAAGGCAATCATGTTATA

FIGURE 210

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76401

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30480, pI: 4.60, NX(S/T): 1

MEAAPSRFMFLLFLLTCELAEEVAAEVEKSSDGPAAQEPTWLTDVPAAMEFIAATEVAVIG
FFQDLEIPAVPILHSMVQKFPGVSGISTDSEVLTHYNITGNTICLFRLLVDNEQLNLEDEDI
ESIDATKLSRFIEINSLHMTVEYNPVTVIGLFNSVIQIHLLIMNKASPEYEENMHRYQKAA
KLFQKGILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDDEWDTLPTAEVSVEHVQNF
CDGFSLSGKLLKENRESEGKTPKVEL

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 143-162

213/249

FIGURE 211

GGAGAGCCGCGGCTGGGACCGGAGTGGGGAGCGCGGCGTGGAGGTGCCACCCGGCGCGGGTG
 GCGGAGAGATCAGAAGCCTCTTCCCCAAGCCGAGCCAACCTCAGCGGGGACCCGGGCTCAGG
 GACGCGGCGGCGGCGGCGGCGACTGCAGTGGCTGGACG**ATGG**CAGCGTCCGCCGGAGCCGGG
 GCGGTGATTGCAGCCCCAGACAGCCGGCGCTGGCTGTGGTCGGTGCTGGCGGGCGGCGCTTGG
 GCTCTTGACAGCTGGAGTATCAGCCTTGGAAGTATATACGCCAAAAGAAATCTTCGTGGCAA
 ATGGTACACAAGGGAAGCTGACCTGCAAGTTCAAGTCTACTAGTACGACTGGCGGGTTGACC
 TCAGTCTCCTGGAGCTTCCAGCCAGAGGGGGCCGACACTACTGTGTGCTTTTTTCCACTACTC
 CCAAGGGCAAGTGTACCTTGGGAATTATCCACCATTAAAGACAGAATCAGCTGGGCTGGAG
 ACCTTGACAAGAAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACACAATGGCACC
 TATATCTGTGATGTCAAAAACCCCTCCTGACATCGTTGTCCAGCCTGGACACATTAGGCTCTA
 TGTCTAGAAAAAGAGAATTTGCCTGTGTTTCCAGTTTGGGTAGTGGTGGGCATAGTTACTG
 CTGTGGTCCTAGGTCTCACTCTGCTCATCAGCATGATTCTGGCTGTCCTCTATAGAAGGAAA
 AACTCTAAACGGGATTACACTGGCTGCAGTACATCAGAGAGTTTGTCAACAGTTAAGCAGGC
 TCCTCGGAAGTCCCCCTCCGACACTGAGGGTCTTGTAAGAGTCTGCCTTCTGGATCTCACC
 AGGGCCCAGTCATATATGCACAGTTAGACCACTCCGGCGGACATCACAGTGACAAGATTAAC
 AAGTCAGAGTCTGTGGTGTATGCGGATATCCGAAAGAAT**TAA**GAGAATACCTAGAACATATC
 CTCAGCAAGAAACAAAACCAAACCTGGACTCTCGTGCAGAAAATGTAGCCCATTACCACATGT
 AGCCTTGGAGACCCAGGCAAGGACAAGTACACGTGTACTCACAGAGGGAGAGAAAGATGTGT
 ACAAAGGATATGTATAAATATTCTATTTAGTCATCCTGATATGAGGAGCCAGTGTTCATGA
 TGAAAAGATGGTATGATTCTACATATGTACCCATTGTCTTGCTGTTTTTGTACTTTCTTTTC
 AGGTCATTTACAATTGGGAGATTTTCAGAAACATTCCCTTTCACCATCATTTAGAAATGGTTTG
 CCTTAATGGAGACAATAGCAGATCCTGTAGTATTTCCAGTAGACATGGCCTTTTAATCTAAG
 GGCTTAAGACTGATTAGTCTTAGCATTTACTGTAGTTGGAGGATGGAGATGCTATGATGGAA
 GCATACCCAGGGTGGCCTTTAGCACAGTATCAGTACCATTATTTGTCTGCCGCTTTTAAAA
 AATACCCATTGGCTATGCCACTTGAAAACAATTTGAGAAGTTTTTTTGAAGTTTTTCTCACT
 AAAATATGGGGCAATTGTTAGCCTTACATGTTGTGTAGACTTACTTTAAGTTTGCACCCTTG
 AAATGTGTCATATCAATTTCTGGATTCATAATAGCAAGATTAGCAAAGGATAAATGCCGAAG
 GTCACTTCATTCTGGACACAGTTGGATCAATACTGATTAAGTAGAAAATCCAAGCTTTGCTT
 GAGAACTTTTGTAACGTGGAGAGTAAAAAGTATCGGTTTTA

FIGURE 212

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76510

><subunit 1 of 1, 269 aa, 1 stop

><MW: 29082, pI: 9.02, NX(S/T): 3

MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKS
TSTTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDDKSDASINIEN
MQFIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVGIVTAVVLGLTLLISMI
LAVLYRRKNSKRDTGCTSESLSPVKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSG
GHHSDKINKSESVVYADIRKN

Signal peptide:

amino acids 1-37

Transmembrane domain:

amino acids 161-183

FIGURE 213

GCCGGCTGTGCAGAGACGCC**ATG**TACCGGCTCCTGTTCAGCAGTGAAGTCCCCGGGGCTGCCGCC
CCCCGGGGGCTTGGCCTCAAGCTGCGGACGACGCGGGGTCCATCAGCGCGCCGGGGCTGCCGCC
TCTCGGCCACGGCTGGGTGCGGGGGCCTCGGGCTGGGGCTGGGGCTGGCGCTCGGGGTGAAGC
TGGCAGGTGGGGCTGAGGGGCGCGGCCCGGCGCAGTCCCCCGCGGCCCGGACCTGAGGCG
TCGCCTCTGGCCGAGCCGCCACAGGAGCAGTCCCTCGCCCCGTGGTCTCCGCAGACCCCGGC
GCCGCCCTGCTCCAGGTGCTTCGCCAGAGCCATCGAGAGCAGCCGCGACCTGCTGCACAGGA
TCAAGGATGAGGTGGGCGCACCGGGGCATAGTGGTTGGAGTTTCTGTAGATGGAAAAGAAGTC
TGGTCAGAAGGTTTAGGTTATGCTGATGTTGAGAACCGTGTACCATGTAAACCAGAGACAGT
TATGCGAATTGCTAGCATCAGCAAAAGTCTCACCATGGTTGCTCTTGCCAAATTGTGGGAAG
CAGGGAAACTGGATCTTGATATTCCAGTACAACATTATGTTCCCGAATTCACAGAAAAAGAA
TATGAAGGTGAAAAGGTTTCTGTGACACAAGATTACTGATTTCCCATTTAAGTGGAAATTCG
TCATTATGAAAAGGACATAAAAAAGGTGAAAGAAGAGAAAGCTTATAAAGCCTTGAAGATGA
TGAAAGAGAATGTTGCATTTGAGCAAGAAAAAGAAGGCAAAAGTAATGAAAAGAATGATTTT
ACTAAATTTAAAACAGAGCAGGAGAATGAAGCCAAATGCCGGAATTCAAAACCTGGCAAGAA
AAAGAATGATTTTGAACAAGGCGAATTATATTTGAGAGAAAAGTTTGAAAATTCAATTGAAT
CCCTAAGATTATTTAAAAATGATCCTTTGTTCTTCAAACCTGGTAGTCAGTTTTTGTATTCA
ACTTTTGGCTATACCCTACTGGCAGCCATAGTAGAGAGAGCTTCAGGATGTAAATATTTGGA
CTATATGCAGAAAATATTCCATGACTTGGATATGCTGACGACTGTGCAGGAAGAAAACGAGC
CAGTGATTTACAATAGAGCAAGG**TAA**ATGAATACCTTCTGCTGTGTCTAGCTATATCGCATC
TTAACACTATTTTATTAATTAAGTCAAATTTTCTTTGTTTCCATTCCAAAATCAACCTGC
CACATTTTGGGAGCTTTTCTACATGTCTGTTTTCTCATCTGTAAAGTGAAGGAAGTAAACA
TGTTTATAAAGTAAAAAAA

FIGURE 214

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76522
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41221, pI: 8.54, NX(S/T): 0
MYRLLSAVTARAAAPGGLASSCGRRGVHQRAGLPPLGHGWVGGLGLGLGLALGVKLAGGLRG
AAPAQSPAAPDPEASPLAEPPEQESLAPWSPQTPAPPSCRCFARAIESSRDLLHRIKDEVGA
PGIVVGVSVDGKEVWSEGLGYADVENRVPCPETVMRIASISKSLTMVALAKLWEAGKLDLD
IPVQHYVPEFPEKEYEGEKVSVTTRLLISHLSGIRHYEKDIKKVKEEKAYKALKMMKENVAF
EQEKEGKSNEKNDFTKFKTEQENEAKCRNSKPGKKKNDFEQGELYLREKFENSIESLRLFKN
DPLFFKPGSQFLYSTFGYTLLAAIVERASGCKYLDYMQKIFHDLDMMLTTVQEENEPVIYNRAR
```

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 39-60

FIGURE 215

GTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTAAGCTCGGAATTCGGCTCG
AGGCTGGTGGGAAGAAGCCGAG**ATG**GCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTCCTGC
TCTTGCTGATGGCGGTAGCAGCGCCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCCGGGACT
GGTGCGCGAGGGGCTGGGGCGGAAGGTCGAGAGGGCGAGGCCTGTGGCACGGTGGGGCTGCT
GCTGGAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGCTCACTGCTCT
GGAACCAGCAGGATGGTACCTTGTCCCTGTCACAGCGGCAGCTCAGCGAGGAGGAGCGGGGC
CGACTCCGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATCCCAAGGCGACCCGG
GGCCCTGGATGGCCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGG
TGGAGTCGCACCTGTCGGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGC
GTGTCGGTGGTGACGCACCCCGGGGGCTGCCGGGGCCATGAGGTGGAGGACGTGGACCTGGA
GCTGTTCAACACCTCGGTGCAGCTGCAGCCGCCACCACAGCCCCAGGCCCTGAGACGGCGG
CCTTCATTGAGCGCCTGGAGATGGAACAGGCCCAGAAGGCCAAGAACCCCCAGGAGCAGAAG
TCCTTCTTCGCCAAATACTGGATGTACATCATTCCCGTCGTCCTGTTCCCTCATGATGTCAGG
AGCGCCAGACACCGGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCC
TTTGCTGTGTGCCACCCTCCCTG**TAA**GTCTATTTAAAAACATCGACGATACATTGAAATGTG
TGAACGTTTTGAAAAGCTACAGCTTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGG
TCCTGATGTACAAGCTTGATTGAAATTCAGTGTGCTCACTTGATACGTTATTCAGAAACCCAAG
GAATGGCTGTCCCCATCCTCATGTGGCTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTAT
TAAACTGTCCCCCAGATCGACACGCAAAAAAAAA

FIGURE 216

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76529

><subunit 1 of 1, 269 aa, 1 stop

><MW: 28004, pI: 5.80, NX(S/T): 1

MAAASAGATRLLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSFEI
DDSANFRKRGSLLWNQQDGTLSLSQRQLSEEEERGRLRDVAALNGLYRVRI PRRPGALDGLEA
GGYVSSFVPACSLVESHLSQDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLELFNTSVQ
LQPPTTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFAKYWMYIIPVVLFLMMSGAPDTGGQ
GGGGGGGGGGGSGLCVPPSL

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 226-243

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 217

GGAGCGCTGCTGGAACCCGAGCCGAGCCGGAGCCACAGCGGGGAGGGTGGCCTGGCGGCCCT
GGAGCCGGACGTGTCCGGGGCGTCCCCGCAGACCGGGGGCAGCAGGTCTGCCGGGGGCCACC
ATGCTGGTGACTGCCTACCTTGCTTTTGTAGGCCTCCTGGCCTCCTGCCTGGGGCTGGA
GTCAAGATGCCGGGCTAAACCCCTGGAAGGGCCTGCAGCAATCCCTCCTTCCTTCGGTTTC
AACTGGACTTCTATCAGGTCTACTTCCTGGCCCTGGCAGCTGATTGGCTTCAGGCCCCCTAC
CTCTATAAACTCTACCAGCATTACTACTTCCTGGAAGGTCAAATTGCCATCCTCTATGTCTG
TGGCCTTGCCTCTACAGTCCTCTTTGGCCTAGTGGCCTCCTCCCTTGTGGATTGGCTGGGTC
GCAAGAATTCTTGTGTCCTCTTCTCCCTGACTTACTCACTATGCTGCTTAACCAA
CAAGACTACTTTGTGCTGCTAGTGGGGCGAGCACTTGGTGGGCTGTCCACAGCCCTGCTCTT
CTCAGCCTTCGAGGCCTGGTATATCCATGAGCACGTGGAACGGCATGACTTCCCTGCTGAGT
GGATCCCAGCTACCTTTGCTCGAGCTGCCTTCTGGAACCATGTGCTGGCTGTAGTGGCAGGT
GTGGCAGCTGAGGCTGTAGCCAGCTGGATAGGGCTGGGGCCTGTAGCGCCCTTTGTGGCTGC
CATCCCTCTCCTGGCTCTGGCAGGGGCCTTGGCCCTTCGAAACTGGGGGGAGA
GGCAGCGTGCCTTCTCAAGGACCTGTGCTGGAGGCCTGCGCTGCCTCCTGT
GTGCTGCTGCTGGGCACCATAACAAGCTCTATTTGAGAGTGTATCTTCATCTTTGTCTTCCT
CTGGACACCTGTGCTGGACCCACACGGGGCCCCCTCTGGGCATTATCTTCTCCAGCTTCATGG
CAGCCAGCCTGCTTGGCTCTTCCCTGTACCGTATCGCCACCTCCAAGAGGTACCACCTTCAG
CCCATGCACCTGCTGTCCCTTGCTGTGCTCATCGTCGTCTTCTCTCTCTTCATGTTGACTTT
CTCTACCAGCCCAGGCCAGGAGAGTCCGGTGGAGTCCTTCATAGCCTTTCTACTTATTGAGT
TGGCTTGTGGATTATACTTTCCCAGCATGAGCTTCCTACGGAGAAAGGTGATCCCTGAGACA
GAGCAGGCTGGTGTACTCAACTGGTTCGGGTACCTCTGCACTCACTGGCTTGCCTAGGGCT
CCTTGTCCCTCCATGACAGTGATCGAAAAACAGGCACTCGGAATATGTTGAGCATTGCTCTG
CTGTCATGGTGATGGCTCTGCTGGCAGTGGTGGGACTCTTCACCGTGGTAAGGCATGATGCT
GAGCTGCGGGTACCTTCACCTACTGAGGAGCCCTATGCCCCTGAGCTG**TAA**CCCCACTCCAG
GACAAGATAGCTGGGACAGACTCTTGAATTCCAGCTATCCGGGATTGTACAGATCTCTCTGT
GACTGACTTTGTGACTGTCCCTGTGGTTTCTCCTGCCATTGCTTTGTGTTTGGGAGGACATGA
TGGGGGTGATGGACTGGAAAGAAGGTGCCAAAAGTTCCTCTGTGTTACTCCCATTTAGAAA
ATAAACACTTTTAAATGATCAAAAAAAAAAAAAA

FIGURE 218

MLVTAYLAFVGLLASCLGLELSRCRAKPPGRACSNPSFLRFQLDFYQVYFLALAADWLQAPY
LYKLYQHYYFLEGQIAILYVCGLASTVLFGLVASSLVDWLGRKNSCVLFSLTYSLCCLTKLS
QDYFVLLVGRALGGLSTALLFSAFEAWYIHEHVERHDFPAEWIPATFARAAFWNHVLAVVAG
VAAEAVASWIGLGPVAPFVAAIPLLALAGALALRNWGENYDRQRAFSRTCAGGLRCLLSDRR
VLLLGTIQALFESVIFIFVFLWTPVLDPHGAPLGII FSSFMAASLLGSSLYRIATSKRYHLQ
PMHLLSLAVLIVVFSLFMLTFSTSPGQESPVESFIAFLLLIELACGLYFPSMSFLRRKVIPET
EQAGVLNWFVRVPLHSLACLGLLVLHDSRKTGTRNMF SICSAVMVMALLAVVGLFTTVVRHDA
ELRVPSPTTEEPYAPEL

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 41-55, 75-94, 127-143, 191-213, 249-270, 278-299,
314-330, 343-359, 379-394, 410-430

FIGURE 219

GCGACGCGCGGCGGGGCGGCGAGAGGAAACGCGGCGCCGGGCGGGGCCCGGCCCTGGAG**ATG**
 GTCCCCGGCGCCGCGGGCTGGTGTGTCTCGTGCTCTGGCTCCCCGCGTGCGTCGCGGCCCA
 CGGCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCCTGGGGACATTCGATACA
 TCTTCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTCACACAAGGTATGAGCAGATT
 CACCTTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACTCAGCAACGGTTTCTTCATCCA
 GGACCAGATTGCTCTGGTGGAGAGGGGGGGCTGCTCCTTCCTCTCCAAGACTCGGGTGGTCC
 AGGAGCACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACAATGACAGCTTCTAC
 GTGGAGATGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTCTTCCTGCTCGG
 CCGAGACGGCTACATGATCCGCCGCTCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTT
 CCATCCCAGTCAATGTCACCAGCATCCCCACCTTTGAGCTGCTGCAACCGCCCTGGACCTTC
 TGG**TAGA**AAGAGTTTGTCCCACATTCCAGCCATAAGTGACTCTGAGCTGGGAAGGGGAAACCC
 AGGAATTTTGCTACTTGGAATTTGGAGATAGCATCTGGGGACAAGTGGAGCCAGGTAGAGGA
 AAAGGGTTTGGGCGTTGCTAGGCTGAAAGGGAAGCCACACCACTGGCCTTCCCTTCCCCAGG
 GCCCCAAGGGTGTCTCATGCTACAAGAAGAGGCAAGAGACAGGCCCCAGGGCTTCTGGCTA
 GAACCCGAAACAAAAGGAGCTGAAGGCAGGTGGCCTGAGAGCCATCTGTGACCTGTCACACT
 CACCTGGCTCCAGCCTCCCCTACCCAGGGTCTCTGCACAGTGACCTTCACAGCAGTTGTTGG
 AGTGGTTTAAAGAGCTGGTGTTTGGGGACTCAATAAACCCCTCACTGACTTTTTAGCAATAAA
 GCTTCTCATCAGGGTTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 220

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76532

><subunit 1 of 1, 188 aa, 1 stop

><MW: 21042, pI: 5.36, NX(S/T): 2

MVPGAAGWCCLVLWLPACVAAHGFRIHDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQ
IHLVPAEPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIISDNAVDNDSF
YVEMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFELLQPPWTFW

Signal peptide:

amino acids 1-20

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 221

TCTGCCTCCACTGCTCTGTGCTGGGATCATGGAACTTGCACTGCTGTGTGGGCTGGTGGTGA
TGGCTGGTGTGATTCCAATCCAGGGCGGGATCCTGAACCTGAACAAGATGGTCAAGCAAGTG
ACTGGGAAAATGCCCATCCTCTCCTACTGGCCCTACGGCTGTCACTGCGGACTAGGTGGCAG
AGGCCAACCCAAAGATGCCACGGACTGGTGCTGCCAGACCCATGACTGCTGCTATGACCACC
TGAAGACCCAGGGGTGCGGCATCTACAAGGACAACAACAAAAGCAGCATACATTGTATGGAT
TTATCTCAACGCTATTGTTTAATGGCTGTGTTTAATGTGATCTATCTGGAAAATGAGGACTC
CGAATAAAAAGCTATTACTAWTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 222

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76538
><subunit 1 of 1, 116 aa, 1 stop
><MW: 12910, pI: 6.41, NX(S/T): 1
MELALLCGLVVMAGVPIQGGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDW
CCQTHDCCYDHLKTQGCGLIYKDNNKSSIHCMDSLQRYCLMAVFNVIIYLENEDSE

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 1-24

N-glycosylation site.

amino acids 86-89

N-myristoylation sites.

amino acids 20-25, 45-50

Phospholipase A2 histidine active site.

amino acids 63-70

Downloaded from www.nrjonline.com by guest on 06/11/16

FIGURE 223

CTCGCTTCTTCCTTCTGGATGGGGGCCCAGGGGGCCCAGGAGAGTATAAAGGCGATGTGGAG
GGTGCCCGGCACAACCAGACGCCAGTCACAGGCGAGAGCCCTGGG**ATG**CACCGGCCAGAGG
CCATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAGATGTAT
GGCCCTGGAGGAGGCAAGTATTTTCAGCACCACTGAAGACTACGACCATGAAATCACAGGGCT
GCGGGTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCCAGGTGAAACTTGGAGACTCCTGGG
ACGTGAAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGGCGAATAC
ATCACAAAAGTCTTTGTGCGCCTTCCAAGCTTTCCTCCGGGGTATGGTCATGTACACCAGCAA
GGACCGCTATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCTGCCTACCCCAGCCAAG
AGGGGCAGGTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTTGGCATCAAGAGCATTGGC
TTTGAATGGAATTATCCACTAGAGGAGCCGACCACTGAGCCACCAGTTAATCTCACATACTC
AGCAAACTCACCCGTGGGTGCG**TAG**GGTGGGGTATGGGGCCATCCGAGCTGAGGCCATCTGT
GTGGTGGTGGCTGATGGTACTGGAGTAACTGAGTCGGGACGCTGAATCTGAATCCACCAATA
AATAAAGCTTCTGCAGAAAA

FIGURE 224

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19600, pI: 5.89, NX(S/T): 1
MHRPEAMLLLLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLVKSQVK
LGDSWDVKLGALGGNTQEVTLQPGYITKVFAFQAFLRGMVMTSKDRYFYFGKLDGQISS
AYPSQEGQVLVGIYGQYQLLGIKSIGFEWNYPLEEPTTEPPVNLTYSANSPVGR
```

Signal peptide:

amino acids 1-22

GCTGAGCGTGTCGCGGGGTACGGGGCTCTCCTGCCCTTCTGGGCTCCAACGCAAGCTCTGTGGCT
GAACTGGGTGCTCATCACGGGAACTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAG
CCCCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTT
TTTTTTTTTAACCGCCCCCTCCCCACCCCCCAAAAAAAGTGTAAAGATGCAAAAACGTAATAT
CCATGAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATT
TATTTGTTCTTGAGAGTGTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCT
CCCAAGGGGTCCAATTTTTCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTC
ACAGGGGCTGTCATGCAACTGGCCCCTAAGCCAAAGCAAAGACCTAAGGACGACCTTTGAA
CAATACAAAGGATGCGGTTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTT
ATAGCCCCCACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTG
TAGGTGTGAAGGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCTCAAGTA
TATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAAGTAAAGTATAAT
CAATTTAAAGGGCTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATAT
TGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCGAATAGAA
TCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATTTACGGAACCTTGGATCTGTCC
TATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTGCAGGAGCTGCTGAGTTT
ACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACC
TGGAACCTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGC
ATGATCAGACTCAAAGAAGCTTCACCTGGAGCACAAATCAATTTTCCAAGCTCAACCTGGCCCT
TTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGTCTATAG
GACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATC
GAAGCTTTTCAGTGGACCCAGTGTTTTTCCAGTGTTGTCCCAGTCTGCAGCGCCTCAACCTGGA
TTCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACA
TCAGTCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGTGGCTG
AAAAGTTTTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGG
AGTAAATGTGATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGT
TTGATCTGGCCAGGGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCAT
GAGAGCAAACCCCTTTGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGC
TGACGCCGAGCACATCTCTTTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCG
TGCTCGTCATCCTGCTGGTTATCTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAG
CTGACGACGCGCTCCCTCATGCGAAGGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAAT
GACTCCAGCAGCCAGGAATTTATGTAGATTATAAACCCACCAACACGGAGACCAGCGAGA
TGCTGCTGAATGGGACGGGACCCTGCACCTATAACAAATCGGGCTCCAGGGAGTGTGAGGTA
TGAACCATTGTGATAAAAAGAGCTCTTAAAGCTGGGAAATAAGTGGTGCTTTATTGAATC
TGGTGACTATCAAGGGAACGCGATGCCCCCCTCCCTTCCCTCTCCCTCTCACCTTTGGTGG
CAAGATCCTTCCTTGTCCGTTTTAGTGCATTATAATACTGGTCATTTTCTCTCATACATA
ATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTTGAAGTCCGGTTTAATATAA
TACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCGATTGTGTTTTAAGATAAACT
TCTTTCATAGGTAAAAA

FIGURE 226

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77301
><subunit 1 of 1, 513 aa, 1 stop
><MW: 58266, pI: 9.84, NX(S/T): 4
MGFNVIRLLSGSAVALVIAPTIVLLTMLSSAERGCPKGCRCGKMVYCESQKLQEIPSSISAG
CLGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYF
LNNTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELL
DLGYNRIRSLARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQNLQWNLKISVIGQTM
SWTWSSLQRLDLGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIGQEILDSWISLNDISLA
GNIWECSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLA
RALPKPTFKPKLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVI
LLVIYVSWKRYPASMKQLQQRSLMRRHRKKKRQSLKQMPSTQEFYVDYKPTNTETSEMLLN
GTGPCTYNKSGSRECEV

Important features of the protein:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 227

AGTTCTGAGAAAGAAGGAAATAAACACAGGCACCAAACCACTATCCTAAGTTGACTGTCCTT
 TAAAT**ATG**TCAAGATCCAGACTTTTCAGTGTACCTCAGCGATCTCAACGATAGGGATCTTG
 TGTTTGCCGCTATTCCAGTTGGTGCTCTCGGACCTACCATGCCAAGAAGATGAAATGTGTGT
 AAATTATAATGACCAACACCCTAATGGCTGGTATATCTGGATCCTCCTGCTGCTGGTTTTGG
 TGGCAGCTCTTCTCTGTGGAGCTGTGGTCCTCTGCCTCCAGTGCTGGCTGAGGAGACCCCGA
 ATTGATTCTCACAGGCGCACCATGGCAGTTTTTTGCTGTTGGAGACTTGGACTCTATTTATGG
 GACAGAAGCAGCTGTGAGTCCAACTGTTGGAATTCACCTTCAAACCTCAAACCCCTGACCTAT
 ATCCTGTTCCCTGCTCCATGTTTTGGCCCTTTAGGCTCCCCACCTCCATATGAAGAAATTGTA
 AAAACAACCT**TGA**TTTTAGGTGTGGATTATCAATTTAAAGTATTAACGACATCTGTAATTCCA
 AAACATCAAATTTAGGAATAGTTATTTTCAGTTGTTGGAAATGTCCAGAGATCTATTCATATA
 GTCTGAGGAAGGACAATTCGACAAAAGAATGGATGTTGGAAAAAATTTTGGTCATGGAGATG
 TTTAAATAGTAAAGTAGCAGGCTTTTGATGTGTCACTGCTGTATCATACTTTTATGCTACAC
 AACCAAATTAATGCTTCTCCACTAGTATCCAAACAGGCAACAATTAGGTGCTGGAAGTAGTT
 TCCATCACATTTAGGACTCCACTGCAGTATACAGCACACCATTTTCTGCTTTAAACTCTTTC
 CTAGCATGGGGTCCATAAAAATTATTATAATTTAACAATAGCCCAAGCCGAGAATCCAACAT
 GTCCAGAACCAGAACCAGAAAGATAGTATTTGAATGAAGGTGAGGGGAGAGAGTAGGAAAAA
 GAAAAGTTTGGAGTTGAAGGGTAAAGGATAAATGAAGAGGAAAAGGAAAAGATTACAAGTCT
 CAGCAAAAACAAGAGGTTTTATGCCCCAACCTGAAGAGGAAGAAATTGTAGATAGAAGGTGA
 AGGAGATTGCTGAAGATATAGAGCACATATAATGCCAACACGGGGAGAAAAGAAAATTTCCC
 CTTTTACAGTAATGAATGTGGCCTCCATAGTCCATAGTGTCTCTGGAGCCTCAGGGCTTG
 GCATTTATTGCAGCATCATGCTAAGAACCTTCGGCATAGGTATCTGTTCCCATGAGGACTGC
 AGAAGTAGCAATGAGACATCTTCAAGTGGCATTTTTGGCAGTGGCCATCAGCAGGGGGACAGA
 CAAAACATCCATCACAGATGACATATGATCTTCAGCTGACAAATTTGTTGAACAAAACAAT
 AAACATCAATAGATATCTAAAAA

FIGURE 228

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77303

><subunit 1 of 1, 146 aa, 1 stop

><MW: 16116, pI: 4.99, NX(S/T): 0

MSRSRLFSVTSAISTIGILCLPLFQLVLSDLPCEEDEMVCVNYNDQHPNGWYIWILLLLVLVA

ALLCGAVVLCLQCWLRRPRIDSHRRTMAVFAVGDLDSIYGTEAAVSPTVGIHLQTQTPDLYP

VPAPCFGPLGSPPPYEEIVKTT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 52-70

FIGURE 229

GAGCGGAGTAAAATCTCCACAAGCTGGGAACAAACCTCGTCCCAACTCCCACCCACCGGCGT
 TTCTCCAGCTCGATCTGGAGGCTGCTTCGCCAGTGTGGGACGCAGCTGACGCCCCGCTTATTA
 GCTCTCGCTGCGTCGCCCCGGCTCAGAAGCTCCGTGGCGGCGGCGACCGTGACGAGAAGCCC
 ACGGCCAGCTCAGTTCTCTTCTACTTTGGGAGAGAGAGAAAAGTCAGATGCCCCTTTTAAACT
 CCCTCTTCAAACTCATCTCCTGGGTGACTGAGTTAATAGAGTGGATACAACCTTGCTGAAG
 ATGAAGAATATACAATATTGAGGATATTTTTTTCTTTTTTTTTTCAAGTCTTGATTTGTGGC
 TTACCTCAAGTTACCATTTTTTCAGTCAAGTCTGTTTGTTGCTTCTTCAGAAATGTTTTTA
 CAATCTCAAGAAAAAATATGTCCCAGAAATTGAGTTTACTGTTGCTTGTATTTGGACTCATT
 TGGGGATTGATGTTACTGCACTATACTTTTCAACAACCAAGACATCAAAGCAGTGTCAGTT
 ACGTGAGCAAATACTAGACTTAAGCAAAAAGATATGTTAAAGCTCTAGCAGAGGAAAATAAGA
 ACACAGTGGATGTGAGAACGGTGCTTCTATGGCAGGATATGCGGATCTGAAAAGAACAATT
 GCTGTCTTCTGGATGACATTTTGCAACGATTGGTGAAGCTGGAGAACAAAGTTGACTATAT
 TGTGTGAATGGCTCAGCAGCCAACACCACCAATGGTACTAGTGGGAATTTGGTGCCAGTAA
 CCACAAATAAAAGAACGAATGTCTCGGGCAGTATCAGATAGCAGTTGAAAATCACCTTGTGC
 TGCTCCATCCACTGTGGATTATATCCTATGGCAGAAAAGCTTTATAATTGCTGGCTTAGGAC
 AGAGCAATACTTTACAATAAAAGCTCTACACATTTTCAAGGAGTATGCTGGATTCATGGAAC
 TCTAATTCTGTACATAAAAAATTTTAAAGTTATTTGTTTGCTTTCAGGCAAGTCTGTTCAATG
 CTGTACTATGTCCTTAAAGAGAATTTGGTAACTTGGTTGATGTGGTAAGCAGATAGGTGAGT
 TTTGTATAAATCTTTTGTGTTTGAGATCAAGCTGAAATGAAAACACTGAAAAACATGGATT
 ATTTCTATAACACATTTATTTAAGTATATAACACGTTTTTTTGGACAAGTGAAGAATGTTTAA
 TCATTCTGTCAATTTGTTCTCAATAGATGTAAGTGTAGACTACGGCTATTTGAAAAATGTG
 CTTATTGTACTATATTTTGTATTCCAATTATGAGCAGAGAAAGGAAATATAATGTTGAAAA
 TAATGTTTTGAAATCATGACCCAAAAGATGTATTGATTTGCACTATCCTTCAGAATAACTGA
 AGGTTAATTATTGTATATTTTTTAAAAATTACACTTATAAGAGTATAATCTTGAAATGGGTAG
 CAGCCACTGTCCATTACCTATCGTAAACATTGGGGCAATTTAATAACAGCATTAAAAATAGTT
 GTAAACTCTAATCTTATACTTATTGAAGAATAAAAGATATTTTTTATGATGAGAGTAACAATA
 AAGTATTCATGATTTTTTCACATACATGAATGTTTCAATTTAAAAGTTTAAATCCTTTGAGTGTCT
 ATGCTATCAGGAAAGCACATTATTTCCATATTTGGGTAAATTTTGCTTTTATTATATTGGTC
 TAGGAGGAAGGGACTTTGGAGAATGGAACCTTGAGGACTTTAGCCAGGTGTATATAATAAA
 GGTACTTTTGTGCTGCATTAAATTGCTTGGAAGTGTTAACATTATATTATATAAGAGTATC
 CTTTATGAAATTTTGAATTTGTATAACAGATGCATTAGATATTCATTTTATATAATGGCCAC
 TTAATAAAGAACATTTAAATATAAACTATGAAGATTGACTATCTTTTCAGGAAAAAAGCT
 GTATATAGCACAGGGAACCTAATCTTGGGTAAATCTAGTATAAAACAAATTATACTTTTAT
 TTAAATTTCCCTTGTAAGCAATCTAATTGCCACATGGTGCCCTATATTTTCATAGTATTTATT
 CTCTATAGTAAGTCTTAAGTGCAGCTAGCTTCTAGATTTAGACTATATAGAATTTAGATAT
 TGTATTGTTTCGTCATTATAATATGCTACCACATGTAGCAATAATTACAATATTTTATTAATA
 TAAATATGTGAAATATTGTTTCATGAAAGACAGATTTCCAAATCTCTCTCTCTCTCTGTA
 CTGTCTACCTTTATGTGAAGAAATTAATTATATGCCATTGCCAGGT

FIGURE 230

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77648
><subunit 1 of 1, 140 aa, 1 stop
><MW: 15668, pI: 10.14, NX(S/T): 5
MFFTISRKNMSQKLSLLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQILDLSKRYVKALAE
ENKNTVDVENGASMAGYADLKRTIAVLLDDILQRLVKLENKVDYIVVNGSAANTTNGTSGNL
VPVTTNKRTNVSGSIR
```

Important features of the protein:

Signal peptide:

amino acids 1-26

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 231

CGCGGCCGGGCCGCCGGGGTGAAGCTGCCGAGGCGGCTGTGGCGCAGGCTTCCAGCCCCAC
CATGCCGTGGCCCCCTGCTGCTGCTGCTGGCCGTGAGTGGGGCCCAGACAACCCGGCCATGCT
TCCCCGGGTGCCAATGCGAGGTGGAGACCTTCGGCCTTTTCGACAGCTTCAGCCTGACTCGG
GTGGATTGTAGCGGCCTGGGCCCCCACATCATGCCGGTGCCCATCCCTCTGGACACAGCCCA
CTTGGACCTGTCTCCAACCGGCTGGAGATGGTGAATGAGTCGGTGTTGGCGGGGCCGGGCT
ACACGACGTTGGCTGGCCTGGATCTCAGCCACAACCTGCTCACCAGCATCTCACCCTACTGCC
TTCTCCCGCCTTCGCTACCTGGAGTCGCTTGACCTCAGCCACAATGGCCTGACAGCCCTGCC
AGCCGAGAGCTTCACCAGCTCACCCTGAGCGACGTGAACCTTAGCCACAACCAGCTCCGGG
AGGTCTCAGTGTCTGCCTTCACGACGCACAGTCAGGGCCGGGCACTACACGTGGACCTCTCC
CACAACCTCATTACCCGCTCGTGCCCCACCCACGAGGGCCGGCCTGCCTGCGCCCAACAT
TCAGAGCCTGAACCTGGCCTGGAACCGGCTCCATGCCGTGCCAACCTCCGAGACTTGCCCC
TGCGCTACCTGAGCCTGGATGGGAACCTCTAGCTGTCAATTGGTCCGGGTGCCTTCGCGGGG
CTGGGAGGCCTTACACACCTGTCTCTGGCCAGCCTGCAGAGGCTCCCTGAGCTGGCGCCAG
TGGCTTCCGTGAGCTACCGGGCCTGCAGGTCTTGACCTGTGCGGCAACCCCAAGCTTAACCT
GGGCAGGAGCTGAGGTGTTTTAGGCCTGAGCTCCCTGCAGGAGCTGGACCTTTCGGGCACC
AACCTGGTGCCCCCTGCCTGAGGCGCTGCTCCTCCACCTCCCGGCACTGCAGAGCGTCAGCGT
GGGCCAGGATGTGCGGTGCCGGCGCCTGGTGCGGGAGGGCACCTACCCCGGAGGCCTGGCT
CCAGCCCCAAGGTGCCCTGCACTGCGTAGACACCCGGGAATCTGCTGCCAGGGGCCCCACC
ATCTTG**TGA**CAAATGGTGTGGCCCAGGGCCACATAACAGACTGCTGTCTGGGCTGCCTCAG
GTCCCGAGTAACCTTATGTTCAATGTGCCAACACCAGTGGGGAGCCCGCAGGCCTATGTGGCA
GCGTCACCACAGGAGTTGTGGGCCTAGGAGAGGCTTTGGACCTGGGAGCCACACCTAGGAGC
AAAGTCTCACCCCTTTGTCTACGTTGCTTCCCCAAACCATGAGCAGAGGGACTTCGATGCCA
AACCAGACTCGGGTCCCCCTCCTGCTTCCCTTCCCCACTTATCCCCCAAGTGCCTTCCCTCAT
GCCTGGGCGGCCTGACCCGCAATGGGCAGAGGGTGGGTGGGACCCCTGCTGCAGGGCAGA
GTTTCAAGTCCACTGGGCTGAGTGTCCCCTTGGGCCCATGGCCAGTCACTCAGGGGCGAGTT
TCTTTTCTAACATAGCCCTTTCTTTGCCATGAGGCCATGAGGCCCGCTTCATCCTTTTCTAT
TTCCCTAGAACCTTAATGGTAGAAGGAATTGCAAAGAATCAAGTCCACCCTTCTCATGTGAC
AGATGGGGAAACTGAGGCCTTGAGAAGGAAAAAGGCTAATCTAAGTTCCTGCGGGCAGTGGC
ATGACTGGAGCACAGCCTCCTGCCTCCCGAGCCCGACCCAATGCACTTTCTTGTCTCCTCTA
ATAAGCCCCACCCCTCCCCGCCTGGGCTCCCCCTTGCTGCCCTTGCTGTTCCCCATTAGCACA
GGAGTAGCAGCAGCAGGACAGGCAAGAGCCTCACAAGTGGGACTCTGGGCCTCTGACCAGCT
GTGCGGCATGGGCTAAGTCACTCTGCCCTTCGGAGCCTCTGGAAGCTTAGGGCACATTGGTT
CCAGCCTAGCCAGTTTCTCACCCCTGGGTGGGGTCCCCCAGCATCCAGACTGGAAACCTACC
CATTTTCCCCCTGAGCATCCTCTAGATGCTGCCCAAGGAGTTGCTGCAGTTCTGGAGCCTCA
TCTGGCTGGGATCTCCAAGGGCCTCCTGGATTCACTCCCCACTGGCCCTGAGCACGACAGC
CCTTCTTACCCTCCCAGGAATGCCGTGAAAGGAGACAAGGTCTGCCCGACCCATGTCTATGC
TCTACCCCCAGGGCAGCATCTCAGCTTCCGAACCCCTGGGCTGTTTCTTAGTCTTCATTTTA
TAAAAGTTGTTGCCTTTTTTAACGGAGTGTCACCTTTCAACCGGCCTCCCCTACCCCTGCTGGC
CGGGGATGGAGACATGTCATTTGTAAAAGCAGAAAAAGGTTGCATTTGTTCACTTTTGTAAT
ATTGTCCTGGGCCTGTGTTGGGGTGTGGGGGAAGCTGGGCATCAGTGGCCACATGGGCATC
AGGGGCTGGCCCCACAGAGACCCACAGGGCAGTGAGCTCTGTCTTCCCCCACCTGCCTAGC
CCATCATCTATCTAACCGGTCTTGATTTAATAAACACTATAAAAGGTTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 232

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77652
><subunit 1 of 1, 353 aa, 1 stop
><MW: 37847, pI: 6.80, NX(S/T): 2
MPWPLLLLLLAVSGAQTTTRPCFPGCQCEVETFGLFDSFSLTRVDCSGLGPHIMPVPIPLDTAH
LDLSSNRLEMVNESVLAGPGYTTLAGLDLSHNLLTSISPTAFSRLRYLESLDLSHNGLTALP
AESFTSSPLSDVNLSHNQLREVSVSAFTTHSQGRALHVDLSHNLIHRLVPHPTRAGLPAPTI
QSLNLAWNRLHAVPNLRDLPLRYLSLDGNPLAVIGPGAFAAGLGGLTHLSLASLQRLPELAPS
GFRELPGQLQVLDLSGNPKLNWAGAEVFSGLSSLQELDLSGTNLVPLPEALLLHLPALQSVSV
GQDVRCRRLVREGTYPRRPGSSPKVPLHCVDTRESAARGPTIL
```

Signal peptide:

amino acids 1-16

Transmembrane domains:

amino acids 215-232, 287-304

FIGURE 233

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGG
GGACCAGAAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATCCCCA
AAACAAGTTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCT
GTTCCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGC
CACGACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCA
GGGAGGCTTGGCAGTTTTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCC
TCTAGTCTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGG
ACTGAAGACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATG
GATTTTCTGAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTA
AGGAGGACTGAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTT
GCTAAGACTCTATCTGGACAGGGTATTTAAAAACTACCAGACCCCTGACCATTATACTCTCC
GGAAGATCAGCAGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCAT
GCCCACATGACATGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCA
CTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTC
TGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGT
CAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACT
GTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGAT
TGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTT
TCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTT
ATTTTTTTTACTTGACATGAACTTTAAAAAAATTACAGATTATATTTATAACCTGACTAG
AGCAGGTGATGTATTTTTTATACAGTAAAAAAAAAAAAACCTTGTAATTCTAGAAGAGTGGCT
AGGGGGGTTATTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGA
TATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATT
GCACATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTAT
CTTCCAGCCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATA
CCAAAAAAAAAAAAAAAAAAAAA

FIGURE 234

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500

><subunit 1 of 1, 261 aa, 1 stop

><MW: 29667, pI: 8.76, NX(S/T): 0

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNGATTGATNSHSDSELRP
EIFSSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNL
QEIRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTP
DHYTLRKISSLANSLFTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALG
ELDILLQWMEETE

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 235

CCGTTATCGTCTTGCGCTACTGCTGA**ATG**TCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCG
CTGACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGC
CGAGCTAGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAG
CTCTTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACA
GCCCTAGGGATCATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCAT
TTACAGACACGTAGTGTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTG
TGTTTGGCAAAAGTGAAGATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATG
GCTGGTGTTATTGGCCAGTTTTTAGCCAATCCAACCTGACCTAGTGAAGGTTTCAGATGCAAAT
GGAAGGAAAAAGGAAACTGGAAGGAAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTG
CAAAAATCTTAGCTGAAGGAGGAATACGAGGGCTTTGGGCAGGCTGGGTACCCAATATACAA
AGAGCAGCACTGGTGAATATGGGAGATTTAACCCTTATGATACAGTGAAACACTACTTGGT
ATTGAATACACCACTTGAGGACAATATCATGACTCACGGTTTATCAAGTTTATGTTCTGGAC
TGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCATCAAAGCAGAATAATGAATCAACCA
CGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCGACTGACTGCTTGATTCAGGCTGT
TCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACCATCTTGGCTGAGAATGACCC
CTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGATGAGTGGAGTCAGTCCA
TTT**TAA**

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77568

><subunit 1 of 1, 323 aa, 1 stop

><MW: 36064, pI: 9.33, NX(S/T): 1

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGA
RESAPYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEH
YPLWKSIVIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFKILAEGGI
RGLWAGWVPNIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTP
ADVIKSRIMNQPRDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLT
YEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 237

CGGACGCGTGGGCGCGGGACGCCGGCAGGGTTGTGGCGCAGCAGTCTCCTTCCTGCGCGCGC
GCCTGAAGTCGGCGTGGGCGTTTGAGGAAGCTGGGATACAGCATTTAATGAAAAATTTATGC
TTAAGAAGTAAAA**ATG**GCAGGCTTCCTAGATAATTTTCGTTGGCCAGAATGTGAATGTATTG
ACTGGAGTGAGAGAAGAAATGCTGTGGCATCTGTTGTGCGAGGTATATTGTTTTTTACAGGC
TGGTGGATAATGATTGATGCAGCTGTGGTGTATCCTAAGCCAGAACAGTTGAACCATGCCTT
TCACACATGTGGTGTATTTTCCACATTGGCTTTCTTCATGATAAATGCTGTATCCAATGCTC
AGGTGAGAGGTGATAGCTATGAAAGCGGCTGTTTAGGAAGAACAGGTGCTCGAGTTTGGCTT
TTCATTGGTTTCATGTTGATGTTTGGGTCACCTATTGCTTCCATGTGGATTCTTTTTTGGTGC
ATATGTTACCCAAAATACTGATGTTTATCCGGGACTAGCTGTGTTTTTTCAAATGCACTTA
TATTTTTTAGCACTCTGATCTACAAATTTGGAAGAACCGAAGAGCTATGGACCT**GA**GATCAC
TTCTTAAGTCACATTTTCTTTTGTATATTCTGTTTGTAGATAGGTTTTTTATCTCTCAGT
ACACATTGCCAAATGGAGTAGATTGTACATTAAATGTTTTGTTTCTTTACATTTTTATGTTT
TGAGTTTTGAAATAGTTTTATGAAATTTCTTTATTTTTTCATTGCATAGACTGTTAATATGTA
TATAATACAAGACTATATGAATTGGATAATGAGTATCAGTTTTTTATTCTGAGATTTAGAA
CTTGATCTACTCCCTGAGCCAGGGTTACATCATCTTGTCATTTTAGAAGTAACCACTCTTGT
CTCTCTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGG
CCGATTGCTTGAGGTCAAGTGTTTGAGACCAGCCTGGCCAACATGGCGAAACCCCATCTACT
AAAAATACAAAATTAGCCAGGCATGGTGGTGGGTGCCTGTAATCCCAGCTACCTGGGAGGC
TGAGGCAGGAGAATCGCTTGAACCCGGGGGGCAGAGGTTGCAGTGAGCTGAGTTTGCGCCAC
TGCACTCTAGCCTGGGGGAGAAAGTGAACTCCCTCTCAAAAAAAGACCACTCTCAGTATC
TCTGATTTCTGAAGATGTACAAAAAATATAGCTTCATATATCTGGAATGAGCACTGAGCCA
TAAAAGGTTTTCAGCAAGTTGTAACCTATTTTGGCCTAAAAATGAGGTTTTTTTTGGTAAAGA
AAAAATATTTGTTCTTATGTATTGAAGAAGTGTACTTTTATATAATGATTTTTTAAATGCCC
AAAGGACTAGTTTGAAAGCTTCTTTTAAAAAGAATTCCTCTAATATGACTTTATGTGAGAA

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 238

MAGFLDNFRWPECECIDWSERRNAVASVVGILFFTGWWMIDAADVYPKPEQLNHAFHTCG
VFSTLAFFMINAVSNAQVRGDSYESGCLGRTGARVWLFIGFMLMFGSLIASMWILFGAYVTQ
NTDVYPGLAVFFQNALIFFSTLIYKFG RTEELWT

Important features:**Signal peptide:**

amino acids 1-44

Transmembrane domains:

amino acids 23-42 (type II), 60-80, 97-117, 128-148

Downloaded from www.sciencedirect.com

FIGURE 239

GTTGATGGCAAACCTTCCTCAAAGGAGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGGCCCA
 CTGGCGGGCCCGCAACACTCCGTCTCACCCCTCTGGGCCCCACTGCATCTAGAGGAGGGCCGTCT
 GTGAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTCAGAAGCTGGCCCAGGGTGGT
 GGTCAGCTGGGTGAGGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAG
 GGAAGTGGGAGCCTCGAGCCCTCGGGTGGAAGCTGACCCCAAGCCACCCTTCACCTGGACAG
GATGAGAGTGTCAGGTGTGCTTCGCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGA
 TGTTTATTCTGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCA
 GCCTCGCCCACCAAGGAGATCCAGGTTAAAAAGTACAAGTGTGGCCTCATCAAGCCCTGCCC
 AGCCAATACTTTTGC GTTTTAAATCTGCAGTGGGGCCGCCAACGTCGTGGGCCCTACTATGT
 GCTTTGAAGACCGCATGATCATGAGTCCTGTGAAAAACAATGTGGGCAGAGGCCTAAACATC
 GCCCTGGTGAATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGG
 AGATGTTATGCACCTAGTGAAATTCCTTAAAGAAATTCGGGGGGGTGCACTGGTGCTGGTGG
 CCTCCTACGACGATCCAGGGACCAAATGAACGATGAAAGCAGGAAACTCTTCTCTGACTTG
 GGGAGTTCCTACGCAAAACAACCTGGGCTTCCGGGACAGCTGGGTCTTCATAGGAGCCAAAGA
 CCTCAGGGGTAAAAGCCCCCTTTGAGCAGTTCTTAAAGAACAGCCAGACACAAACAATAACG
 AGGGATGGCCAGAGCTGCTGGAGATGGAGGGCTGCATGCCCCCGAAGCCATTT**TAG**GGTGGC
 TGTGGCTCTTCCTCAGCCAGGGGCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCCG
 GCAGGGGCTGAGGAGGAGGAGCAGGGGGTGTGCGTGGAAGGTGCTGCAGGTCCTTGACGC
 TGTGTGCGGCCTCTCCTCCTCGGAAACAGAACCCTCCCACAGCACATCCTACCCGGAAGACC
 AGCCTCAGAGGGTCCTTCTGGAACCAGCTGTCTGTGGAGAGAATGGGGTGCTTTTCGTAGGG
 ACTGCTGACGGCTGGTCCTGAGGAAGGACAAACTGCCCAGACTTGAGCCCAATTAAATTTTA
 TTTTTTGCTGGTTTTTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 240

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59814

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24963, pI: 9.64, NX(S/T): 1

MRVSGVLRLLALIFAIVTTWMFIRSMSFSMKTIRLPRWLAASPTKEIQVKKYKCGLIKPCP
ANYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAFDMYSG
DVMHLVKFLKEIPGGALVLVASYDDPGTKMNDESRKLFSDLGSSYAKQLGFRDSWVFIGAKD
LRGKSPFEQFLKNSPDTNKYEGWPELLEMEGCMPPKPF

Important features:**Signal peptide:**

amino acids 1-15

ATP/GTP-binding site motif A (P-loop).

amino acids 184-191

N-glycosylation site.

amino acids 107-110

FIGURE 241

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGAT
 CCAGAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACA
 GCAAGAGAAGCAGAGATAAATACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTC
 TCACTCCTCCCTCCCTCTCTCTCTGCTGTCTAGTCCTCTAGTCCTCAAATTCCCAGTCCC
 CTGCACCCCTTCCTGGGACACT**ATG**TTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCC
 TGGCTGCAGATGGGGGTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTTGGCCA
 GCCTCTTACCCTGAGTGTGGAAACAATGCCAGTCGCCCATCGATATTCAGACAGACAGTGT
 GACATTTGACCCTGATTTGCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGC
 CTTTGGACCTGCACAACAATGGCCACACAGTGCAACTCTCTCTGCCCCTCTACCCTGTATCTG
 GGTGGACTTCCCCGAAAATATGTAGCTGCCCAGCTCCACCTGCACTGGGGTCAGAAAGGATC
 CCCAGGGGGGTCAGAACACCAGATCAACAGTGAAGCCACATTTGCAGAGCTCCACATTGTAC
 ATTATGACTCTGATTCCTATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCCTGGCT
 GTCCTGGGCATCCTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTCTGAG
 TCACTTGCATGAAGTCAGGCATAAAGATCAGAAGACCTCAGTGCCTCCCTTCAACCTAAGAG
 AGCTGCTCCCCAACAGCTGGGGCAGTACTTCCGCTACAATGGCTCGCTCACAACCTCCCCCT
 TGCTACCAGAGTGTGCTCTGGACAGTTTTTTATAGAAGGTCCCAGATTTCAATGGAACAGCT
 GGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAAGCTTCTGGTACAGA
 ACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTTCATCCAAGCAGGA
 TCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGCTGTCTCTG
 CCTTCTCCTGGCTGTTTTATTTCAATTGCTAGAAAGATTTCGGAAGAAGAGGCTGGAAAACCGAA
 AGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCAT**TAA**ATTCCTTCTCAGATACCA
 TGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTG
 GCCAGAAACACTGTAGGAGTAGTAAGCAGATGTCTCCTTCCCCTGGACATCTCTTAGAGAG
 GAATGGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTA
 GGAGGAAATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACTCTGTTTAGTTGCAGGG
 GAAGTTTGGGATATACCCCAAAGTCCTCTACCCCTCACTTTTATGGCCCTTTCCTAGATA
 TACTGCGGGATCTCTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATA
 TATTTGGAAATTAAAGTTTCTGACTTT

FIGURE 242

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62812

><subunit 1 of 1, 337 aa, 1 stop

><MW: 37668, pI: 6.27, NX(S/T): 1

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPAASYPECGNNAQSPIDIQTDSVTFDPDLP
ALQPHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQ
INSEATFAELHIVHYSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRH
KDQKTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGT
FSTEEEPSKLLVQNYRALQPLNQRMVFASFQAGSSYTTGEMLSLGVGILVGCLCLLLAVYF
IARKIRKKRLENRKSVVFTSAQATTEA

Important features of the protein:**Signal peptide:**

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 243

AATTTTTACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTC
GTGGACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGG
ATCAACTCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTC
CGGATCAGGGAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTG
ATACCATTAACACAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGG
AATGACACCTGGTACCCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAAC
TGCACCCACATGTGTTACCAATTTTTGTGCACACAACCTTGAGCCCAGGGCACTATCCTAAGC
TCAGAGGAATTGCCACAAATCTTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCAT
CCTGCCCACCAGTCAGGCAGGGGCTAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAG
GAGCAGGTGTAAATCCTGCCACCCAGGGAACCCAGCAGGCCGCCTCCCAACTCCCAGTGGC
ACAGATGACGACTTTGCAGTGACCACCCCTGCAGGCATCCAAAGGAGCACACATGCCATCGA
GGAAGCCACCACAGAATCAGCAAATGGAATTCAGTAAGCTGTTTCAAATTTTTTCAACTAAG
CTGCCTCGAATTTGGTGATACATGTGAATCTTTATCATTTGATTATATTATGGAATAGATTGA
GACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATTTACCTGAAAATATTCTTGAAATT
TCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAACAATAATTCAATGGATAAAT
CTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATATTAAAACATATTTGGAAA
ACTGGAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 244

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHHPVLPFI FVTQLGAQG TILSSEE
LPQIFTSLIIHSLFP GGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

247/249

FIGURE 245

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCG
GAGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCG
GCAGCCGGGAGCC**ATG**CGACCCCAGGGCCCCGCCGCCTCCCCGCAGCGGCTCCGCGGCCTCC
TGCTGCTCCTGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAG
CAAAAGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGG
GCCAGCAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTG
GGATCCCAGGTTCGGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAG
GAGTCCTGGACACCCAACTACAAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCT
TGGGAAAATTGCGGAGTGACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGT
TCAGTGGCTCACTTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTTACATTC
AATGGAGCTGAATGTTTCAGGACCTCTTCCCATTGAAGCTATAATTTATTTGGACCAAGGAAG
CCCTGAAATGAATTCAACAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAG
GAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGGTTGGCACTTGTTTCAGATTACCCAAAA
GGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAA**TA**
AATGCTTTAATTTTTCATTTGCTACCTCTTTTTTTTATTATGCCTTGGAATGGTTCACTTAAAT
GACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGA
CCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTGCTTCAATCAAAAGT
GGTTTCAATATTTTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTA
TAATTTGGAATATGTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTAAAAAATA
TAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGTAAAT
AAAAATTATTTCCAACA

FIGURE 246

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393

><subunit 1 of 1, 243 aa, 1 stop

><MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPA
PGRDGSPGANVIPGTPGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSSLNYGIDLKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPEMN
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250